A DETAILED INVESTIGATION INTO THE ENZYMATIC SPECIFICITY OF PORCINE PEPSIN

By

CHETANA MANOHAR RAO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1994

This dissertation is dedicated to my parents, Manohar and Vijaya, for instilling upon me the value of education at an early age

ACKNOWLEDGMENTS

My undergraduate teachers were instrumental in moulding my interests in biochemistry in general and the study of enzymes in particular. While serendipity brought me to the University of Florida, in retrospect, I am glad I came here. The University provided me with an environment conducive to my intellectual growth, scientific maturity, and the freedom to pursue my academic interests.

Active research on aspartic proteinases in Prof. Ben Dunn's laboratory gave me the opportunity to further my understanding of enzymes. It gave me excellent exposure to various members of the family of aspartic proteinases such as pepsin, cathepsin D, cathepsin E, rhizopuspepsin, and HIV protease. This led to my dissertation study on porcine pepsin. A deep appreciation goes to my mentor Prof. Ben M. Dunn for his valuable guidance, untiring support, and patience throughout my graduate career. His positive attitude, even in the face of disappointing data, was very encouraging. I hope to imbibe at least some of his qualities as a researcher. I also thank all my committee members Drs. Brian Cain, Harry Nick, David Silverman, and David Richardson for their valuable suggestions.

While working in this lab, I came in contact with people who helped me solve some of the teething problems that I encountered during my graduate studies; some of them greatly influenced me without being aware that they did. Bill Farmerie extended his expertise to get around stumbling blocks in cloning. Todd Lowther and Paula Scarborough provided stimulating scientific discussions as well as invaluable friendship.

Finally, this endeavor would have been incomplete without the emotional support, love and understanding of my entire family. In particular, I thank my wonderful parents Vijaya and Manohar, my brother Sanjay, my uncle Bala, and last but not least, my husband Prasad whose everlasting patience and understanding made it all worthwhile.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS iii

LIST OF TABLES vii

LIST OF FIGURES ix

ABBREVIATIONS x

ABSTRACT xii

CHAPTERS

1 PROTEOLYTIC ENZYMES 1
  Introduction 1
  Serine Proteinases 2
  Cysteine Proteinases 8
  Metallo Proteinases 10
  Aspartic Proteinases 11
    General Characteristics 13
    Porcine Pepsin 16

2 MATERIALS AND METHODS 28
  Introduction 28
  Materials 28
  Methods 29
    Isolation of Plasmid DNA 29
    Mutagenesis 31
    First PCR Reaction 33
    Purification of PCR Products 34
    Second PCR Reaction 35
    Restriction Digests 36
    CIP Treatment 36
    Ligation 39
    Preparation of Competent Cells 40
    Transformation 41
    Double Stranded DNA Sequencing 42
    Expression of Glu13 and Glu287 Mutants 45
    SDS-PAGE Analysis of Expression of Protein 47
    Isolation of Inclusion Bodies 47
    Refolding 48
    Gel Filtration Chromatography 49
    Activation to Mature Form 49
    Ion Exchange Chromatography 50

N-Terminal Sequencing 50
Spectrophotometric Analyses 51

3 SUBSTRATE SPECIFICITY REQUIREMENTS OF PORCINE PEPSIN: ANALYSES USING CHROMOGENIC OCTAPEPTIDE SUBSTRATES WITH SYSTEMATIC SUBSTITUTIONS IN THE P_5-P_3' POSITIONS 54
Introduction 54
Results 55
Kinetic Analyses 55
pH Dependence Studies 62
Discussion 65

4 ROLE OF GLU13 AND GLU 287 RESIDUES IN GENERAL/SPECIFIC ELECTROSTATIC INTERACTIONS 67
Introduction 67
Results 68
Expression of Porcine Pepsinogen 68
Activation 68
Purification 71
N-Terminal Sequencing 71
Kinetic Analysis 74
Discussion 83
The S_2 Subsite 83
The S_2 Subsite 84

5 COMPARISON OF KINETIC PARAMETERS OF PORCINE PEPSIN WITH CATHEPSIN D AND CATHEPSIN D 89
Introduction 89
Results 90
S_3 Subsite 90
S_2 Subsite 92
Discussion 96

6 CONCLUSIONS AND FUTURE DIRECTIONS 107

LIST OF REFERENCES 110

BIOGRAPHICAL SKETCH 120

LIST OF TABLES

Table page

1-1 Sequence alignment of some of the mammalian aspartic proteinases 15

2-1 Primers synthesized for the XhoI silent mutation 31

2-2 Primers designed for PCR mutagenesis at position 13 of the pepsinogen gene . 32

2-3 Primers designed for PCR mutagenesis at position 287 of the pepsinogen gene 33

2-4 Primers synthesized for sequencing the Glu13 and Glu287 mutations 43

3-1 Kinetic parameters of peptide substrates of the form Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu with substitutions in P_5 to P_3' positions 57

3-2 Kinetic parameters of peptide substrates of the form Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu with systematic substitutions in the P_2 position . 61

3-3 pK_a values of enzyme-catalyzed cleavage of peptide substrates of the form Lys-Pro-Xaa-Yaa-Phe*Nph-Arg-Leu 64

4-1 Kinetic parameters of peptide substrates of the form Lys-Pro-Xaa-Yaa-Phe Nph-Arg-Leu with substitutions in P_3 and P_2 for wild type and wild type recombinant porcine pepsin 75

4-2 Analysis of Glu13, Glu287 and double mutants of porcine pepsin at pH 3 5 with the substrate K(Lys)-P(Pro)-A(Ala)-K(Lys)-F(Phe)*X(Nph)-R(Arg)-L(Leu) with systematic substitutions in the P_2 and P_3 positions 76

4-3 pKa values of enzyme-catalyzed cleavage of peptide substrates of the form Lys-Pro-Xaa-Yaa-Phe*Nph-Arg-Leu 80

4-4 K_i values with a competitive inhibitor LP-130 82

5-1. Specificity constants (k_{cat}/K_m) for the enzyme-catalyzed cleavage of peptide substrates of the form Lys-Pro-Xaa-Yaa-Phe*Nph-Arg-Leu . . . 91

5-2. Specificty constants (k_{cat}/K_m) of peptide substrates of the form Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu with systematic substitutions in the P_2 position for porcine pepsin, human cathepsin D and human cathepsin E . . . 93

5-3 Analysis of the double mutant cycles of porcine pepsin with the cleavage of the peptide of the form Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu with substitutions in the P_3 position . . . 103

5-4. Analysis of the double mutant cycles of porcine pepsin with the cleavage of the peptide of the form Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu with substitutions in the P_2 position . . . 104

LIST OF FIGURES

Figure | page

1-1 Superposition of the C_α-carbon backbone traces of porcine pepsin, human cathepsin E and human cathepsin D . . . 14

1-2 A ribbon trace of the backbone of porcine pepsin . . . 17

1-3 Top view of the active site explaining the Schechter and Berger nomenclature . . . 19

1-4 A top view of the active site of porcine pepsin complexed with the Abbott inhibitor A63218 . . . 24

2-1 Plasmid map of pET-3b with the pepsinogen gene insert . . . 37

2-2 Site-directed mutagenesis by overlap extension . . . 38

3-1 pH dependence curve of the peptide Lys-Pro-Arg-Lys-Phe*Nph-Arg-Leu with wild type porcine pepsin . . . 63

4-1 SDS-PAGE analysis of the expression of recombinant mutant porcine pepsinogen . . . 69

4-2 SDS-PAGE analysis of the activation of Glu287Gln mutant pepsinogen . . . 70

4-3 Gel filtration chromatography elution profile of Glu13Ala mutant enzyme . . . 72

4-4 Anion exchange chromatography elution profile of Glu13Ala mutant enzyme . . . 73

ABBREVIATIONS

amp	ampicillin
ATP	adenosine triphosphate
AU	absorption units
BPN'	*Bacillus amyloliquefaciens N*
BSA	bovine serum albumin
C	carboxy
CAPS	3-(cyclohexylamino)-1-propanesulfonic acid
cm	centimeters
CTP	cytosine triphosphate
$CuSO_4$	copper sulfate
d	deoxy
dd	dideoxy
DMSO	dimethylsulfoxide
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
gm(s)	gram(s)
GTP	guanosine triphospahte
HCl	hydrochloric acid
HPLC	High Performance Liquid Chromatography
I	ionic strength
IPTG	isopropylthio-β-D-galactopyranoside
K_2HPO_4	potassium phosphate dibasic
kcal	kilocalories
k_{cat}	turnover number
KCl	potassium chloride
K_i	inhibition constant
K_m	Michaelis-Menten constant
Lb	Luria Broth
MES	2-(4- morpholino)-ethane sulfonic acid

mg	milligram
$MgCl_2$	magnesium chloride
min	minutes
ml(s)	milliliter(s)
mM	millimolar
MOPS	3-(N-morpholino) propane-sulfonic acid
N	amino
Na_2 EDTA	di-sodium ethylenediaminetetraacetic acid
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanograms
NH_4Cl	ammonium chloride
nM	nanomolar
nm	nanometers
Nph	*p*-nitrophenylalanine
NTP	nucleotide triphospahte
OD	optical density
ori	origin
PAGE	polyacrylamide gel electrophoresis
por	porcine
PPGN	porcine pepsinogen
PVDF	polyvinylidene difluoride
r. p. m	revolutions per minute
rec	recombinant
s	second
SDS	sodium dodecyl sulfate
β-ME	β-mercaptoethanol
tet	tetracycline
Tricine	N-[Tris-(hydroxymethyl) Methyl] glycine
Tris	tris (hydroxymethyl) aminomethane
TTP	thymine triphospahte
v/v	volume/volume
w/v	weight/volume
[S]	substrate concentration
μg(s)	microgram(s)
μl(s)	microliter(s)

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

A DETAILED INVESTIGATION INTO THE ENZYMATIC SPECIFICITY OF PORCINE PEPSIN

By

Chetana Manohar Rao

August, 1994

Chairman Dr. Ben M. Dunn
Major Department: Biochemistry and Molecular Biology

Pepsin, a member of the aspartic proteinase family, which is involved in the digestion of proteins in the stomach, has been a valuable model system for investigation of enzymatic specificity. The active site cleft of porcine pepsin is made of distinct subsites, S_5-S_3', all of which may contribute to the specificity of the enzyme. The work presented here is a multifaceted study to evaluate the specific structure-activity relationship of porcine pepsin.

Kinetic studies utilizing chromogenic octapeptide substrates, at increasing pH values, were indicative of general and/or specific cation binding preferences in the S_3 and the S_2 subsite of porcine pepsin. The examination of the crystal structure suggested that Glu13 and the Glu287 residues in the S_3 and S_2 subsites, respectively, may be

involved in these electrostatic interactions. Residue 287 within the S_2 region of the active site cleft is variable among the aspartic proteinase family, with Glu, Gln, Met, and Gly occurring in various enzymes. An alanine or a glutamine is found at position 13 in the other members of this family.

A site-directed mutagenesis approach combined with a high resolution crystal structure and in-depth kinetic analyses were undertaken to verify the role of the two residues in specific and/or general electrostatic interactions in the enzyme. The Glu13 was replaced with an alanine or a glutamine. The residue at position 287 was substituted with a methionine or a glutamine. A double mutant, with a glutamine at position 13 and a methionine at position 287, was also generated. The wild type and the mutant proteins were overexpressed in *E. coli* and subsequently purified to homogeneity. All the mutant enzymes were kinetically characterized.

The results have established the origin of specific and/or general electrostatic interactions in the S_3 and S_2 subsite of porcine pepsin. Also significant are the changes in binding requirements and catalysis of the mutant enzymes altering a single residue namely, Glu287, in the S_2 subsite

CHAPTER 1
PROTEOLYTIC ENZYMES

Introduction

Protein-ligand interactions play an integral role in living systems. Enzymatic catalysis is a common example of such an interaction. The encounter between a substrate and an enzyme is controlled by an intricate interplay of a variety of forces which include hydrogen bonding, hydrophobic and electrostatic interactions. The contribution of each of these forces is necessary to understand structure-function relationships.

Advances over the past decade in understanding the structure of proteins have led to attempts to correlate the structure-activity relationship of enzymes. A number of enzymes have been investigated in depth and these studies have revealed the intricacies of enzymatic specificity and catalysis at the molecular level.

Examination of kinetic specificity, in conjunction with detailed three-dimensional structural analyses, has led to the identification of the geometry and the components of the active site of the family of proteinases. Based on these characteristics and the mechanism of action, the proteinases have been categorized into 4 classes, namely,

1. Serine proteinases
2. Cysteine proteinases
3. Metallo proteinases
4. Aspartic proteinases

Proteinases have an extended binding cleft made up of distinct subsites each of which contributes to the enzymatic specificity (Schechter and Berger, 1967). The interactions between the amino acid side chains of the substrate and the side chains within the subsites of the enzyme in the active site of proteolytic enzymes provide a basis for substrate preferences.

The ability to alter a protein by changing the specific amino acids using protein engineering has helped us understand the structure-function relationships in proteins. Several enzymes have been explored through protein engineering to evaluate the role of non-covalent interactions in catalysis and to understand the role of molecular recognition in protein function. The studies focusing on the importance of molecular interactions taking place in specific subsites critical in enzymatic specificity for some of the enzymes will be discussed in this chapter.

Serine Proteinases

Two classes of serine proteinases, which include 1) the mammalian serine proteinases e.g. chymotrypsin, trypsin, elastase, and 2) bacterial serine proteinases, an example of which is subtilisin, have been extensively studied. These enzymes are similar in terms of their mechanism. The active site is composed of a His-Asp-Ser catalytic triad. Catalysis proceeds via a tetrahedral transition state intermediate via acylation and deacylation steps on the serine (Kraut, 1977). However, they differ in their specificity towards substrates. Trypsin cleaves at lysine or arginine. Chymotrypsin prefers large hydrophobic side chains such as phenylalanine, tyrosine and tryptophan. Elastase likes small hydrophobic residues such as alanine (Fersht, 1977). Subtilisin, the bacterial serine proteinases, also prefers hydrophobic residues in the P_1 position of the peptide. The origin of the differences in specificity lie in the S_1 binding pocket of the three enzymes. The size of the pocket and the nature of the amino acids that define this pocket confer the specificity of the individual enzymes. Both the classes have a similar active site geometry, although they differ in the overall amino acid sequences.

Among the bacterial serine proteinases, subtilisin has been extensively studied. To provide a background for studies to be described in Chapters 3, 4 and 5, I will briefly summarize some related work on this enzyme. Protein engineering has been used with the cloned subtilisin gene by various groups to create an enzyme with greater oxidative stability (Estell et al., 1985), to determine the role of hydrogen bonding in the stabilization of the transition state (Bryan et al., 1986), to introduce a disulfide bond into

---

subtilisin (Wells and Powers, 1986) and to change the pH activity profile of the enzyme (Thomas et al., 1985).

The three-dimensional crystal structures of subtilisin complexed with a number of inhibitors has been solved (Bode et al., 1987; McPhalen and James, 1988; Takeuchi et al., 1991a, 1991b). Based on this structural information, it was suggested that there are two types of interactions important in enzymatic specificity, namely the hydrogen bonds between the backbone of the enzyme and the inhibitor and the specific hydrophobic interactions between the side chains of the P_1 and P_4 positions of the peptides which interact with the S_1 and S_4 subsites in the enzyme, respectively (Bech et al., 1993).

Estell et al. explored the steric and the hydrophobic effects on enzymatic specificity of subtilisin (1986). To analyze the contribution of these forces they generated 12 noncharged mutations in the S_1 subsite. Based on the three-dimensional structure, the S_1 binding pocket is defined by the following residues: Leu126, Pro129, Ala152, Glu156 and Gly166. The glycine at position 166 is highly conserved among the bacterial serine proteinases. Mutations were introduced at this position. The results were indicative of changes in the enzymatic specificity toward the nature of the amino acid in the P_1 position of the peptide substrate. When a hydrophobic residue was placed at position 166, enhanced (16-fold) k_{cat}/K_m values were exhibited for peptide substrates with a hydrophobic residue in the P_1 position. Substitutions of larger groups at both position 166 and the P_1 position of the substrate decreased the catalytic efficiency. This decrease was proposed to result due to steric hindrance.

These studies were followed by designing the substrate specificity of subtilisin by engineering electrostatic interactions (Wells et al., 1987). The objective was to try to convert the S_1 subsite specificity of subtilisin to that of the mammalian serine proteinase trypsin. As mentioned earlier, trypsin prefers to cleave substrates with a positively charged residue in the P_1 position of the substrate. Glu156 and Gly166 were replaced systematically with 16 different amino acids. Double mutants at both positions were also analyzed. The charged substitutions in the enzyme showed improved catalytic efficiency toward the oppositely charged P_1 substitutions (up to 1900-fold). Decreased efficiency was noted toward similarly charged substrates.

Gron et al. (1992), in an attempt to understand the binding requirements of subtilisin BPN', had reported the strong preference for aromatic residues at the P_4 position of the peptide substrate. This effect was suggested to take place due to the presence of Tyr104 in the S_4 subsite of subtilisin BPN'. The aromatic ring of Tyr104 positions itself parallel to the aromatic ring of the substrate leading to a productive interaction (Kraut, 1977). Subtilisin 309, however, has a valine at position 104 instead of tyrosine. This enzyme also exhibited a strong preference for aromatic groups in the P_4 position of the peptide. To evaluate the role of position 104 on enzymatic specificity, valine in subtilisin 309 was replaced with seven different amino acids (Bech et al., 1992). Steric, hydrophobic and electrostatic effects in the S_4 binding pocket were explored. Substituting an arginine at position 104 resulted in an 18-fold increase in k_{cat}/K_m values toward a P_4 aspartate substrate in comparison to the wild type enzyme. This result was proposed to take place due to a substrate induced conformational change so that the arginine side chain in the enzyme can make a productive interaction with the aspartate side chain in the substrate. This was in agreement with the mobility shift of the tyrosine residue in subtilisin BPN' (Takeuchi et al., 1991a). Comparison of the kinetic parameters for the Val104Tyr and Val104Phe mutants revealed identical preferences. The Val104Tyr mutant showed improved binding toward peptides with hydrophobic residues in the P_4 position. No change was observed with hydrophilic substitutions of the substrate. Based on these observations the authors concluded that the S_4/P_4 interactions have a profound influence on the catalytic activity of the enzyme. However, this was not only due to the amino acid occupying position 104. It was probably a combined effect of all the residues that define the S_4 pocket that are involved in the binding of the P_4 residue.

All the studies demonstrated the importance of the S_1 and the S_4 subsite in the enzymatic specificity of subtilisin. Bech et al. (1993) further explored the detailed interactions in the S_4 subsite of subtilisin 309. The conserved nonbulky residues that lie along the rim of the S_4 pocket in subtilisin 309 are Gly102, Ser128, Ser130 and Ser132. A bulky phenylalanine residue was independently introduced at each of these positions. Double mutants at position 104 and 132 were also explored to study the interdependence of the effects of substitutions at this position. Previous studies had indicated a preference for hydrophobic residues in the P_4 position of the substrate (Gron et al., 1992). However, introduction of a bulky phenylalanine residue at position 102 and 128 impaired this effect. According to the authors this was a result of blocking the S_4 binding pocket by the phenylalanine so that the hydrophobic interactions were no longer productive. Supporting this observation, no change in specificity was observed for the hydrophilic substitutions in the P_4 position of the peptide. This indicated that within the S_4 subsite the orientations of the hydrophobic residue in the P_4 position of the peptide differed from the hydrophilic substitutions. The interdependency of the double mutant was also examined (Val104Ala/Ser132Phe). Based on the double mutant cycle, to be discussed in detail in Chapter 5, if the mutations were independent, then the interaction energies of the transition state complex in the double mutant would be the sum of the energies of the single mutants (Carter et al., 1984). Such a simple additivity did not exist in the double mutant cycles. These results were suggestive of a flexible S_4 region. The significance of the nonbulky residues in the S_4 pocket were proposed to reduce steric hindrance and allow the larger hydrophobic residues to bind productively.

All the studies, discussed above, focused on molecular interactions involved in the specific subsites important in enzymatic catalysis. Similar studies have been carried out with the mammalian serine proteinases which will not be discussed here (Craik at al., 1985, 1987; Corey and Craik, 1992; Perona at al., 1993).

Cysteine Proteinases

The cysteine proteinases are a group of enzymes which depend on a thiol group of a cysteine residue for catalytic activity. In mammalian tissues, the lysosomal cathepsins B, H, S and L function in protein degradation (Barrett, 1975, Barrett and McDonald, 1980; Barrett and Kirachka, 1981). The most thoroughly characterized cysteine proteinase is the plant enzyme, papain. The three-dimensional structure of papain has been solved to high resolution (Kamphuis et al., 1984, 1985). The active site consists of a catalytic triad of Cys, His and Asp, and catalysis proceeds via a thiol ester intermediate. Even though the catalytic mechanism has been characterized, very few studies on structure-function relationship have been reported

The nature of the residue at the P_2 position of the substrate has been shown to be significant in defining the specificity of the enzyme papain. The S_2 subsite has a preference for bulky nonpolar side chains, such as phenylalanine or tyrosine, at the P_2 position (Berger and Schechter, 1970). The specificity of the mammalian cysteine proteinase, cathepsin B, is different from that of papain. Cathepsin B tolerates positively charged residues in the P_2 position (Barrett and Kirschke, 1981). This is due to the presence of a glutamate residue in cathepsin B replacing a serine residue in papain (Khouri et al., 1991). Khouri et al. used site-directed mutagenesis as a tool to alter the S_2 subsite specificity of papain to that of the mammalian cathepsin B. Based on the sequence alignments and the three-dimensional structure of papain (Kamphuis et al., 1984), two residues Val133 and Ser205 in the S_2 subsite of papain were replaced with Ala133 and Glu205. The mutant enzyme (Val133Ala/Ser205Glu) exhibited increased cleavage of peptides containing an arginine in the P_2 position by a factor of 22. These results confirmed the interaction of the glutamate residue in the S_2 subsite of cathepsin B and the P_2 arginine residue in the peptide substrate. The mutations produced changes that were designed to introduce cathepsin B-like specificity into papain.

Studies establishing the role of Asp158 and the hydrogen bonding interactions involving the side chain of Asp158 in the catalytic mechanism of papain have also been carried out (Menard et al 1990; 1991)

Other subsites have not been explored in depth using site-directed mutagenesis.

Metallo Proteinases

Metallo proteinases include the exopeptidases carboxypeptidase A and B which have similar structure and active sites. Carboxypeptidase A is well characterized kinetically (Vallee et al 1983). The mechanism of these enzymes is still questionable. The active site components include a zinc residue bound to two glutamates and one histidine residue. Among the metallo endopeptidases thermolysin is a classic example. The active site geometry of thermolysin is closely related to the metallo-carboxypeptidases.

The matrix metalloproteinases (MMPs) play an important role in physiological and pathophysiological roles in matrix degradation (Lyons et al., 1991, Woessner, 1991). All MMPs are made up of three domains, namely, the N-terminal domain, the C-terminal domain and the catalytic domain. The human MMPs are made up of eight different enzymes divided into four classes. The first class includes the fibroblast type (Stricklin et al., 1977; Goldberg et al., 1986), the interstitial collagenases and the neutrophil type collagenases (Hasty et al., 1990; Mallya et al., 1990;

Mookhtiar and van Wart, 1990). The second category is the gelatinases or the type-IV collagenases (Salo et al., 1985). The third category has the same domain as the interstitial collagenases but does not exhibit collagenase activity. The fourth class is known as the matrilysins (Muller et al., 1988). The N-terminal domain of all MMPs are well conserved. The carboxy domain is highly variable. The substrate specificity is controlled by both the domains.

Protein engineering approach has not been exploited to detail the molecular interactions in the specific subsites of metallo proteinases conferring enzymatic specificity.

Aspartic Proteinases

The aspartic proteinase family of enzymes functions in a wide variety of physiological and pathophysiological processes. Aspartic proteinases are produced in the human body by a variety of cells. Some of these proteins are secreted and exert their effects in the extracellular spaces, examples of which are pepsin, gastricsin and renin. Cathepsin D and cathepsin E, on the other hand, are intracellular enzymes. The aspartic proteinases are of considerable importance in the control of various biological processes in mammalian systems. Renin, produced in the kidneys, catalyzes the conversion of angiotensinogen to angiotensin I, which is a precursor of angiotensin II, a major factor in the control of blood pressure (Greenlee, 1987). Cathepsin D, a lysosomal enzyme, which functions in the intracellular degradation of proteins, may be involved in the activation of proenzymes or prohormones and also has been shown to be involved in breast tumor invasiveness (Pillai and Zull, 1986; Brouillet et al., 1990). The least characterized of the aspartic proteinases is cathepsin E. It is an intracellular enzyme located in a variety of cells and tissues in the human body such as the gastrointestinal tract, urinary organs and lymphoid tissue (Sakai et al., 1989; Yamamoto et al., 1991). The physiological function of this enzyme is still nebulous. However, Lees et al (1990) have shown that cathepsin E is responsible for the conversion of big endothelin to endothelin, which is a vasoconstrictor more potent than angiotensin II. Pepsin and gastricsin are two aspartic proteinases found in the gastric mucosa. Upon secretion, into the lumen of the stomach, they play an important role in the digestion of proteins. Recently, interest has increased in this family of enzymes due to the discovery of the HIV proteinase which has the signature sequence of Asp-Thr-Gly in the active site common to all aspartic proteinases. This protein catalyzes the cleavage of the *gag-pol* translation product critical in the continuation of the life cycle of the virus (Graves, 1991).

General Characteristics

Aspartic proteinases are characterized by a pH optimum in the acid range and their susceptibility to inhibition by pepstatin, an acetylated pentapeptide isolated from *Streptomyces* (Umezawa et al., 1970). There is a high degree of primary and tertiary structural homology among the enzymes of this family (see Figure 1-1, Table 1-1). The bilobal structure is made up of nearly equivalent N- and C- terminal domains and is dominated by orthogonally packed β-sheets with small α-helical segments (see Figure 1-2). The active site is located at the interface between the two domains (Sali et al., 1992). The active sites of all the members are superimposable with minor variations.

Pepsin, the archetype of this family, functions in the digestive system and has been implicated in the manifestations of gastric ulcers (Walker and Taylor, 1980; Pearson et al., 1986). The active site of pepsin, and all the other aspartic proteinases, is an extended cleft composed of distinct subsites, S_5 -S_3' (Schechter and Berger, 1967), all of which may contribute to the specificity of the enzyme.

A substrate or an inhibitor binds to the active site in an extended β-strand conformation with the side chains spanning into alternating subsites which flank the catalytic site (Hofmann et al., 1984) (See Figure 1-3.)

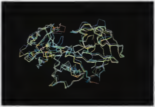

Figure 1-1. Superposition of the C_α-carbon backbone traces of porcine pepsin (orange), human cathepsin E (green) and human cathepsin D (blue).

Table 1-1 Sequence alignment of some of the mammalian aspartic proteinases

```
               1          10         20         3032       40
Porpep[a]   IGDEPLENY  LDTEYFGTIG IGTPAQDFTV IFDTGSSNLW VPSVYCS  S
HuCatD[b]   PIPEVLKNY  MDAQYYGEIG IGTPPQCFTV VFDTGSSNLW VPSIHCKLLD
HuCatE[c]   SAKEPLINY  LDMEYFGTIS IGSPPQNFTV IFDTGSSNLW VPSVYCT  S
HuRenin[d]  TSSVILTNY  MDTQYYGEIG IGTPPQTFKV VFDTGSSNVW VPSSKCSRLY

               50         60         70          80         90
Porpep      LACSDHNQFN PDDSSTFEAT SQELSITYGT G SMTGILGY DTVQV-----
HuCatD      IACWIHHKYN SDKSSTYVKN GTSFDIHYGS G.SLSGYLSQ DTVSVPCQSA
HuCatE      PACKTHSRFQ PSQSSTYSQP GQSFSIQYGT G SLSGIIGA DQVSV-----
HuRenin     TACVYHKLFD ASDSSSYKHN GTELTLRYST G TVSGFLSQ DIITV-----

                         100        110        120        130
Porpep      ------GGIS DTNQIFGLSE TEP GSFLYY APFDGILGLA YPSI SASGA
HuCatD      SSASALGGVK VERQVFGEAT KQP GITFIA AKFDGILGMA YPRI SVNNV
HuCatE      ------EGLT VVGQQFGESV TEP GQTFVD AEFDGILGLG YPSL AVGGV
HuRenin     ------GGIT VT QMFGEVT EMP ALPFML AEFDGVVGMG FIEQ AIGRV

                 140        150        160         170        180
Porpep      TPVFDNLWDQ GLVSQDLFSV YLSSNDDS.  GSVVLLGGID SSYYTGSLNW
HuCatD      LPVFDNLMQQ KLVDQNIFSF YLSRDPDAQP GGELMLGGTD SKYYKGSLSY
HuCatE      TPVFDNMMAQ NLVDLPMFSV YMSSNPEGGA GSELIFGGYD HSHFSGSLNW
HuRenin     TPIFDNIISQ GVLKEDVFSF YYNRNSQS L GGQIVLGGSD PQHYEGNFHY

                    190        200             210  215     220
Porpep      VPVSV EGYW QITLDSITMD GETIA CSGG CQAIVDTGTS LLTGPTSAIA
HuCatD      LNVTR KAYW QVHLDQVEVA S GLTLCKEG CEAIVDTGTS LMVGPVDEVR
HuCatE      VPVTK QAYW QIALDNIQVG GTVMF CSEG CQAIVDTGTS LITGPSDKIK
HuRenin     INLIK TGVW QIQMKGVSVG SSTL LCEDG CLALVDTGAS YISGSTSSIE
```

Table 1-1 continued

	230	240	250	260	270
Porpep	NIQSDIGASE	NSDGEMVIS	CSSIDSLPDI	VFTINGVQYP	LSPSAYIL
HuCatD	ELQKAIGAVP	LIQGEYMIP	CEKVSTLPAI	TLKLGGKGYK	LSPEDYTLKV
HuCatE	QLQNAIGAAP	VDGEYAVE	CANLNVMPDV	TFTINGVPYT	LSPTAYTLLD
HuRenin	KLMEALGAKK	RLF DYVVK	CNEGPTLPDI	SFHLGGKEYT	LTSADYVFQE

	280	290	300	310	320
Porpep	QDDDSCTS	GFEGMDVPTS	SG ELWILGD	VFIRQYYTVF	DRANNKVGLA
HuCatD	SQAGKTLCLS	GFMGMDIPPP	SG PLWILGD	VFIGRYYTVF	DRDNNRVGFA
HuCatE	FVDGMQFCSS	GFQGLDIHPP	AG PLWILGD	VFIRQFYSVF	DRGNNRVGLA
HuRenin	SYSSKKLCTL	AIHAMDIPPP	TG PTWALGA	TFIRKFYTEF	DRRNNRIGFA

	326
Porpep	PVA
HuCatD	EAARL
HuCatE	PAVF
HuRenin	LAR

[a]Porpep = Porcine pepsin, [b]HuCatD = Human cathepsin D, [c]HuCatE = Human cathepsin E, [d]HuRenin = Human Renin The Glu13 and the Glu287 residues in porcine pepsin are underlined

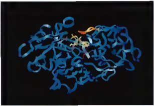

Figure 1-2. A ribbon trace of the backbone of porcine pepsin in blue. Shown in yellow is the Abbott inhibitor A63218 bound in the active site (Abad-Zapatero et al., 1991). Shown in red are the catalytic aspartates 32 and 215 as well as the Glu13 residue in the S_3 subsite and the Glu287 residue in the S_2 subsite of the enzyme. Note the bilobal structure with a pseudo two-fold axis of symmetry composed of nearly equivalent N- and C- terminal domains. Shown in orange is the flap region.

Despite the overall structural similarities among the aspartic proteinase family of enzymes, each proteinase must have subtle distinctions which enables it to fulfill its own function in its own environment. Three-dimensional structures have been elucidated for these enzymes by X-ray crystallography (Bott et al., 1982; Andreeva et al., 1983; James and Sielecki, 1986; Blundell et al., 1987; Cooper et al., 1987; Suguna et al., 1987a; Sali et al., 1989; Abad-Zapatero et al., 1990; Sielecki et al., 1990; Veerapandian et al., 1990; Parris et al., 1991).

The differences in specificity and activity of the various aspartic proteinases may be accounted for by discrete alterations of individual amino acid residues contributing to some or all of the subsites comprising the active site cleft. Such differences have been investigated by kinetic studies employing synthetic substrates and inhibitors (Dunn et al., 1987; Jupp et al., 1990; Scarborough et al., 1991; Kay and Dunn, 1992; Lowther et al., 1991; Rao et al., 1991).

Porcine Pepsin

As mentioned earlier, pepsin is the archetypal member of this family. The chief cells of the gastric mucosa are responsible for the biosynthesis of pepsinogen which is the zymogen form of pepsin (Foltmann, 1981).

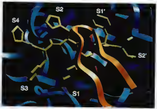

Figure 1-3. Top view of the active site explaining the Schecter and Berger nomenclature (Schechter and Berger, 1967). The Upjohn inhibitor U70531 is shown in yellow. The sequence of the inhibitor is DHis-Pro-Phe-His-Phe-Phe-Val (Suguna et al., 1987b) (code number 3APR). The cleft accommodates at least seven amino acids of a substrate in and extended β-strand conformation. The side chains are shown extending into alternating subsite pockets which flank the catalytic site. Shown in red are the two catalytic aspartates (Asp32 and Asp215).

At acidic pH, pepsinogen undergoes self-activation with the cleavage of a 44 amino acid activation peptide to generate mature pepsin (Herriott, 1939). Kinetic studies on the activation of pepsinogen have revealed that at low pH (<3.0), the conversion from pepsinogen to pepsin is intramolecular (Al-Janabi et al., 1972; Mcphie, 1972; Marciniszyn et al., 1976). The amino acid sequence of porcine pepsinogen has been determined by several groups (Ong and Perlmann, 1968; Pedersen and Foltmann, 1973; Sepulveda et al., 1975). The three-dimensional structure of porcine pepsinogen indicates that the 44 amino acid prosegment makes inhibitory interactions with the pepsin portion of the zymogen (James and Sielecki, 1986). Comparing the crystal structure of pepsinogen and pepsin it has been shown that the conformational changes taking place during the activation process are extensive (Andreeva et al., 1983; Sielecki et al., 1990). The position of the N-terminal Ile residue in active pepsin is about 40 Å distant from its position in pepsinogen (James and Sielecki, 1986).

Pepsin consists of a single polypeptide chain made up of 326 amino acids with a molecular weight of 35,000 daltons. The mature enzyme consists of the N-terminal domain composed of residues 1-175 and a C-terminal domain with residues 176-326 (Andreeva et al., 1983).

The crystal structures of native and inhibitor-complexed porcine pepsin has been resolved to 1.8 Å which provide information about the protein/ligand interactions (Andreeva et al., 1983; Sielecki et al., 1990). Abad-Zapatero et al (1990) have described the structure of porcine pepsin complexed with two Abbott Pharmaceuticals inhibitors, A63218 and A66702. Both the inhibitors are bound to the active site in an extended conformation occupying the S_4 to S_1' subsites. Conformational changes were observed in two regions of pepsin upon binding of the inhibitor. the flap region (residues 72-81) and a portion of the C-terminal end. These studies have proven useful in designing specific inhibitors to improve potency and rationalize the structural changes taking place in the enzyme on ligand binding. Such results can then be correlated to the kinetic data observed using synthetic peptides.

The catalytic efficiency of pepsin has been shown to be dependent upon the length of the peptide and nature of the amino acid side chains that make up the active site cleft initially referred to as "secondary interactions" by Fruton (Fruton, 1970; Sachdev and Fruton, 1970). Initial studies of these secondary interactions utilized di- and tri-peptides, which exhibited poor binding and cleavage rates. Further studies included increasing the length of the peptide to eight residues which showed a remarkable increase in binding and cleavage (Dunn et al., 1984; Blum et al., 1985; Hofmann et al., 1988). These studies gave an insight into the length of the peptides that could optimally fit into the active site of porcine pepsin.

There is also a preference for large hydrophobic aromatic residues on either side of the scissile bond of aspartic proteinases (Dunn, 1992). Using this binding mode, a series of chromogenic octapeptide substrates were synthesized and tested with a wide variety of aspartic proteinases of animal and microbial origin. Results from these studies reflected the hydrophobic character of the S_3 subsite among the mammalian enzymes (Dunn et al., 1986).

Following these studies, cleavage of a variety of peptides tested with three members of the family was examined over a range of pH values (Dunn et al., 1987). The changes in kinetic parameters as a function of pH revealed information about the ionization of functional groups in the ES (enzyme-substrate) complex, free enzyme or free substrate. This study provided evidence for strong and specific hydrogen bonding interactions in the S_2 and S_3 subsites of porcine pepsin.

Following these observations, the cation binding specificity in the subsites of porcine pepsin was examined using octa- and undeca-peptide substrates. Lysine and arginine were placed in seven different positions and the pH dependent effects on the kinetic parameters were studied (Pohl and Dunn, 1988). Two hypotheses were put forward to explain the pH dependent increase in the k_{cat}/K_m values of these positively charged peptides. They were as follows: 1) specific electrostatic interactions exist in the active site cleft involving ion pairs between an enzyme and a ligand, and

2) ligand binding is directed by general electrostatic interactions between the positive charges on the substrates and the negative charges present on the surface of the protein which are predominant in porcine pepsin. This type of interaction was referred to as "non-specific electrostatic binding." Inspection of the crystal structure of porcine pepsin suggested that the residues that may be involved in the specific electrostatic interactions were Glu13 in the S_3 subsite and Glu287 in the S_2 subsite (see Figure 1-4)

Kuzmic et al. (1991) continued the above work to differentiate between the two hypotheses. They utilized positively charged statine-based pepsin inhibitors to carry out their study, at increasing pH values. Based on their results they proposed that hydrophobic attractions rather than electrostatic forces were predominant in the S_2 and S_3 subsites because all the charged inhibitors used in the study bound weakly compared to their neutral analogues. They also suggested that the trend observed in Pohl's work (i.e. increase in k_{cat}/K_m with increasing pH for polycationic substrates) was due to nonspecific binding. Kuzmic et al. (1991) in their studies with the positively charged inhibitors revealed nonspecific binding which was dependent on the charges on the enzyme and the ligand and on the ionic strength of the medium.

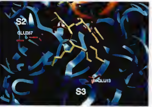

Figure 1-4. A top view of the active site of porcine pepsin (blue) complexed with the Abbott inhibitor A63218 (yellow) (Abad-Zapatero et al., 1991). Shown in red are the two catalytic aspartates (Asp 32 and Asp215). Also shown in red are the two residues Glu13 in the S_3 subsite and Glu287 in the S_2 subsite of the enzyme proposed to be involved in specific and/or general electrostatic interactions.

He inferred that the binding of a ligand to porcine pepsin takes place in two steps: fast nonspecific electrostatic association, followed by secondary interactions, which are predominantly hydrophobic, taking place in the active site cleft.

These studies addressed the specific interactions taking place in the active site of the enzyme, as well as interactions due to the overall charge on the protein at the pH and ionic strength conditions used in the enzyme assay. However, it is important to note that the studies by Kuzmic et al. (1991) were done with statine-based inhibitors, which may not accurately detail specific subsite interactions. This is because the lysine in the P_2 position of the inhibitor could interact with the S_1' subsite of the enzyme rather than the S_2 subsite. This interaction is due to the presence of the statine residue in the P_1-P_1' position leaving the S_1' subsite unoccupied.

Site-directed mutagenesis is a very effective tool to explore the role of specific amino acids that are important in the activity and specificity of an enzyme. Lin et al. (1992) used this approach to study the effect of substituting the active site residue (Asp32Ala) in pepsinogen on the autocatalytic conversion of pepsinogen to pepsin. The difference observed in the autoactivation of the mutant and the wild type indicated that the intramolecular activation of pepsinogen takes place in the active site where the catalytic aspartates are located because the active site mutant

(Asp32Ala) did not cleave pepsinogen to mature pepsin. Chymosin, a mammalian aspartic proteinase used in the cheese industry, has been a model enzyme for mutagenesis studies. Strop et al. (1990) combined site-directed mutagenesis and X-ray crystallography to evaluate the role of Val111 in chymosin which borders the S_1 and S_3 subsites. Changing the valine to phenylalanine (present in porcine pepsin) changed the size of the S_1 and S_3 subsites which made large residues in the P_1 position of the substrate less acceptable to the mutated enzyme. A similar enzyme has been found in fungal strains and is called *Mucor* rennin. Aikawa et al. (1991) used site-directed mutagenesis in *Mucor* rennin as a tool to understand the role of Tyr75, a residue that is conserved among all aspartic proteinases. The results from the Tyr75Phe mutation indicated that Tyr75 may be involved in both substrate binding and catalysis.

The combination of a kinetic approach using synthetic substrates and inhibitors, X-ray crystallography and site-directed mutagenesis will be an excellent system to understand these specific interactions taking place in the S_2 and S_3 subsites in the active site cleft of porcine pepsin.

The present work outlines detailed kinetic studies designed to explore potential interactions in the active site cleft of porcine pepsin using chromogenic octapeptide substrates with two bulky aromatic residues Phe and Nph in the P_1 and P_1' positions, respectively (Schechter and Berger, 1967; Dunn, 1992). The substitutions of a series of mean acids in the P_5-P_3' positions were used to probe the specific interactions such as hydrophobic attractions, van der Waals forces, hydrogen bonding and electrostatic interactions taking place in the individual binding pockets leading to catalysis

In order to evaluate the roles of two residues, Glu13 in the S_1 subsite and Glu287 in the S_2 subsite of the active site cleft of porcine pepsin, site directed mutagenesis was used to generate mutants at Glu13 and Glu287 positions. These mutants were expressed, purified to homogeneity and characterized pH dependence studies were carried out with the wild type as well as the mutant enzymes using synthetic substrates and inhibitors to validate the hypothesis. These studies have enabled us to understand specific molecular interactions taking place in the S_1 and the S_2 pockets of the enzyme and to evaluate the role of Glu13 in the interaction with the substrate at different pH values. It has also given us information into the role of Glu287 in the S_2 subsite of porcine pepsin in enzymatic specificity

CHAPTER 2
MATERIALS AND METHODS

Introduction

Preliminary kinetic characterization of porcine pepsin, utilizing chromogenic octapeptide substrates, to be explained in Chapter 3, has provided information about enzymatic specificity. Based on the three-dimensional crystal structure (Abad-Zapatero et al., 1991) residues were chosen for mutagenesis to support the hypothesis on the binding requirements and catalytic efficiency of porcine pepsin.

In order to test these hypotheses, I have used site directed mutagenesis as a tool to study the effects on changes of the residues on the enzymatic specificity.

This chapter describes 1) the cloning and mutagenesis of porcine pepsin, 2) expression and isolation of inclusion bodies, 3) solubilization and refolding to produce mature enzyme, 4) kinetic characterization of mature enzyme.

Materials

The pET-3b (Studier et al., 1990) expression vector containing the wild type porcine pepsinogen gene was a gift from Prof. Jordan Tang, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma. Synthetic oligonucleotides used in mutagenesis were synthesized at the Interdisciplinary Center for Biotechnology Research DNA Synthesis Core at the University of Florida. All bacterial growth media and ultrapure urea were from GIBCO-Bethesda Research Laboratory. Acrylamide for sequencing and SDS-PAGE were from Bio-Rad. Other reagents were of highest commercial grade quality either from Sigma or Fisher Chemicals. The chromogenic peptide substrates used for kinetic characterization were synthesized by the Protein Chemistry Core Facility, University of Florida, using an Applied Biosystems Peptide 430A synthesizer. All the peptides were verified for purity (≥90%) utilizing reverse phase HPLC. Quantitation of the concentration of the stock solutions of the peptides was performed by the Protein Chemistry Core Facility with the use of a Beckman System 6300 High Performance Amino Acid Analyzer. The N-terminal sequence analysis was also carried out by the Protein Chemistry Core Facility with Applied Biosystems 470A and 473A Protein Sequencers. All recipes to make media were from Molecular Cloning: A Laboratory Manual (Maniatis et al., 1982).

Methods

Isolation of Plasmid DNA

Five milliliters of LB media (950 ml of distilled deionized water, 10 gm of bacto-tryptone, 5 gm of bacto-yeast extract and 10 gm of NaCl, pH adjusted to 7.0 with 5 N NaOH, sterilized) was inoculated with a single colony of *E. coli* HMS174 cells harboring the recombinant pET-3b plasmid with the pepsinogen gene. These inoculates were grown to saturation overnight (12-14 hours). The next day the plasmid DNA was isolated using a Magic™ Miniprep Kit from Promega. Briefly, the 5 ml culture was pelleted into a 1.5 ml sterile microcentrifuge tube and the supernatant was discarded. The pellet was resuspended in 200 µl of cell resuspension solution (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 100 µg/ml RNAse A) and vortexed. To this was added 200 µl of cell Lysis solution (0.2 M NaOH, 1% SDS) and mixed gently by inverting the tubes several times until the solution appeared clear. This was followed by the addition of 200 µl of neutralization solution (2.55 M potassium acetate, pH 4.8) followed by inverting the tubes to ensure proper mixing. The tubes were subjected to centrifugation at 12,000 rpm (11,000 x g) for 10 minutes. The supernatant was transferred to another sterile microfuge tube.

To the supernatant was added 1 ml of Magic™ Miniprep DNA Purification Resin. The purification resin is a proprietary formulation that binds DNA effectively. The tubes were mixed and the contents transferred to a 3 ml syringe. To the end of the syringe was attached a mini-column. The contents were pushed out gently. The mini-column was washed with 3 ml of column wash solution (200 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50% ethanol). The column was removed from the syringe and fitted into another sterile microfuge tube. The tube was centrifuged at 12,000 rpm (11,000 x g) for 20 seconds. The column was then transferred to another sterile tube and the DNA was eluted with 50 µl of sterile water pre-heated to 70°C. The purified DNA was stored at -20°C for further use.

Mutagenesis

To create mutations of the Glu13 and Glu287 residues, oligonucleotides with the specific codon changes were designed and synthesized. In the case of the mutations at the Glu13 position, a unique restriction site Xho I was introduced at the 3' end of the Glu13 codon (70 base pairs 3' of the Glu13 codon) in the pepsinogen gene. The purpose of this exercise was to minimize the region for sequencing for verification of the Glu13 mutants. The oligonucleotides synthesized for the Xho I silent mutation are shown in Table 2-1.

Table 2-1. Primers synthesized for the Xho I silent mutation

Strand	Sequence (5' to 3')	Position
+	CTA TAG GGA GAC CAC AAC GGT TTC C	4538-4562
+	GAG CAA CCT GTG GGT GCC CTC TGT C	4890-4915
-	ACA GGT TGC TCG AGC CGG TGT CAA AGA	4855-4873

Primers synthesized to make changes at the Glu13 position are listed below in Table 2-2.

Table 2-2. Primers designed for PCR mutagenesis at position 13 of the pepsinogen gene.

Strand	Mutation	Sequence (5'-3')	Position
+	Glu13Ala	G AAC TAC CTG GAT ACT GCC TAC TTT GGC AC	4806-4837
+	Glu13Gln	G AAC TAC CTG GAT ACT CAA TAC TTT GGC AC	4806-4837
+	---	CTA TAG GGA GAC CAC AAC GGT TTC C	4538-4562
-	---	GTA TCC AGG TAG TTC TCA AGG GGC TCA TCG	4790-4820
-		GTC TCT CCG TCC ATG GTG ATG CTG T	5344-5369

For the Glu287 mutants, in addition to the codon change at position 287, a unique restriction site (Spe I) was engineered for rapid screening of mutants by restriction digests. The primers used for all the PCR mutagenesis at position 287 and the additional codon change for the Spe I restriction site are listed below in Table 2-3.

Table 2-3 Primers designed for PCR mutagenesis at position 287 of the pepsinogen gene

Strand	Mutation	Sequence (5'-3')	Position
+	Glu287Met	TGC ACT AGT GGC TTC ATG GGC ATG GAT G	5629-5657
+	Glu287Gln	TGC ACT AGT GGC TTC CAG GGC ATG GAT G	5629-5657
+	---	TAC TGG CAG ATT ACC TTG GAC	5350-5371
-	---	AAG CCA CTA GTG CAG CTG TCA TCG T	5618-5643
-	---	ATA GGC GTA TCA CGA GGC	532-550

The plasmid map of the pET-3b with the pepsinogen insert is shown in Figure 2-1

Mutagenesis was carried out using a modified version of the overlap extension method of site directed mutagenesis by PCR (polymerase chain reaction) (Ho et al , 1989).

Figure 2-2 outlines the general procedure used for PCR mutagenesis. PCR was carried out as follows

First PCR Reaction

To the first tube was added 0.5 µl of plasmid DNA (approximately 2 nanograms), 1 µl primer a (100 µM stock solution), 1 µl primer b (100 µM stock solution), 5 µl of 10x Vent polymerase buffer (100 mM KCl, 200 mM Tris-HCl, pH 8.8, 100 mM ammonium sulfate, 20 mM magnesium sulfate, 1% Triton

X-100), 0.5 µl of 20 mM dNTPs, 0.5 µl of 100X acetylated BSA, 40.7 µl of sterile water and 1 µl of Vent polymerase (2 units/µl). In the second tube the same components were added with the exception of primers c and d replacing primers a and b.

PCR was carried out in a Gene Machine II Programmable Thermal Controller (USA Scientific Plastics). The tubes were heated at 94°C for 30 seconds to denature the DNA template. The annealing of the primers to the template was carried out at 60°C for 50 seconds. This was followed by the extension reaction at 72°C for 2 minutes. The cycle of denaturation, annealing and extension was repeated 25 times. The tubes were refrigerated at 4°C.

Purification of PCR Products

Purification of the first PCR reaction products was carried out utilizing agarose gel electrophoresis. The entire contents of the PCR reaction (50 µl) was directly loaded on a 3% Nu-Sieve low melt GTG Agarose gel (FMC BioProducts). The gel was subjected to electrophoresis at 70 Volts until the dye front was 1 cm from the bottom of the gel. The gel was stained with ethidium bromide and visualized on a long wave uv light source (Fotodyne). Bands of the expected size (products AB and CD) were excised, transferred to individual microfuge tubes, and subjected to purification using the Magic™ PCR Miniprep Kit (Promega).

The microfuge tubes containing the individual gel slices were melted at 70°C for 5 minutes. One milliliter of the Magic™ PCR prep resin was added to each of the tubes and vortexed vigorously for 1 minute. The contents were then transferred to a syringe with a mini-column at the end. The contents were pushed out gently. The mini-column was washed with 3 ml of column wash solution (200 mM NaCl, 20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50% ethanol). The column was removed from the syringe and fitted into another sterile microfuge tube and the DNA was eluted with 50 µl of sterile water. The purified PCR product was stored at -20°C for further use.

Second PCR Reaction

The second round of PCR reaction was carried out as follows.

In a tube were mixed 5 µl of product ab and 5 µl of product cd of the first PCR reaction, 1 µl primer a (100 µM stock solution), 1 µl primer d (100 µM stock solution), 5 µl of 10X Vent polymerase buffer, 0.5 µl of 20 mM dNTPs, 0.5 µl of 100X acetylated BSA, 32.0 µl of sterile water and 0.5 µl of Vent polymerase. The tube was subjected to identical PCR conditions as the first PCR reaction. The products were purified on a 1% low melt Nu-Sieve GTG Agarose gel. The correct size band was excised from the gel after staining with ethidium bromide and purified using the Promega Kit.

Restriction Digests

In the case of the Glu13 mutants, 20 µl of the second PCR products (approximately 4 µg) was treated with 3 µl of NEB buffer 2 (10 mM Tris-HCl, 10 mM magnesium chloride, 50 mM NaCl, 1 mM DTT, pH 7.9) 1 µl XbaI (15,000 units/ml) and 1 µl XhoI (20,000 units/ml) and 5 µl of sterile distilled water. For the Glu287 mutants all the contents of the restriction digests were the same with the exception of the enzymes NcoI (1000 units/ml) and EcoRI (10,000 units/ml) with NEB buffer 3 (50 mM Tris-HCl, 10 mM magnesium chloride, 100 mM NaCl, 1 mM DTT, pH 7.9) were used. The tubes were incubated at 37°C for 4 hours. The contents of the tubes were subjected to electrophoresis on a 1% Nu-Sieve GTG agarose gel along with the pET-3b plasmid containing the wild-type pepsinogen insert cut with the respective enzyme pairs namely, XbaI/XhoI or NcoI/EcoRI. The fragments were excised from the gel upon completion of electrophoresis and purified using the Magic™ PCR Miniprep Kit (Promega) as described earlier.

CIP (Calf Intestinal Alkaline Phosphatase) Treatment

Alkaline phosphatases catalyses the removal of 5' phosphate group from DNA, RNA and ribo- and deoxy-ribonucleoside triphosphates. Since CIP-treated fragments lack the 5' phosphate they cannot self ligate.

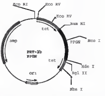

Figure 2 1. Flasmid map of pET 3b with the pepsinogen gene insert

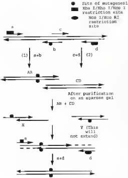
● Sits of mutagenesis.
Xba I/Xho I/Nco I restriction site
Nco I/Eco RI restriction site
a
c
b
d
(1)
a+b
c+d
(2)
AB
CD
After purification on an agarose gel
AB + CD
X
Y (This will not extend)
a
a+d
d
MUTANT PRODUCT

Figure 2-2. Site-directed mutagenesis by overlap extension. Arrows represent the 5' to 3' direction. The mutagenesis sites are denoted by closed circles.

This decreases the vector background in cloning procedures (Maniatis et el., 1982) The CIP treatment was carried out as follows:

To a sterile microfuge tube was added approximately 400 nanograms of plasmid DNA (45 µl), 5 µl of NEB buffer 2 and 1 µl of CIP (1 unit/µl) The tube was incubated for 50°C for 1 hour This was followed by the addition of 0.5 µl of 0.5M EDTA and incubation at 75°C for 15 minutes. Greater than 95% of CIP can be inactivated by this process. In order to ensure complete inactivation of CIP, 1.5 µl of 20% SDS (Sodium dodecyl sulfate) was added followed by 0.5 µl of Proteinase K (20 mg/ml stock solution) The tube was incubated for 20 minutes at 50°C and the entire contents were loaded on a 0.5% Nu-Sieve GTG agarose gel and electrophoresis was carried out. Upon completion, the gel was stained with ethidium bromide and visualized on a long wave uv light source (Fotodyne) Bands of the expected size were excised, transferred to individual microfuge tubes and subjected to purification using the Magic™ PCR Miniprep Kit (Promega).

Ligation

The mutant DNA fragment from the second PCR reaction was ligated into the CIP-treated pET-3b wild type papainogen vector as follows:

In tube 1 approximately 40 nanograms (4 µl) of CIP-treated plasmid was added to approximately 200 nanograms (12 µl) of PCR product, 2 µl of T4 DNA ligase buffer (500 mM Tris-HCl, pH 7.8, 100 mM magnesium chloride, 100 mM DTT, 10 mM ATP), 1 µl 20 mM ATP and 1 µl T4 DNA ligase (400 units/µl).

In tube 2 approximately 40 nanograms (4 µl) of CIP-treated plasmid was added to 12 µl of sterile distilled water, 2 µl of T4 DNA ligase buffer (500 mM Tris-HCl, pH 7.8, 100 mM magnesium chloride, 100 mM DTT, 10 mM ATP), 1 µl 20 mM ATP and 1 µl T4 DNA ligase (400 units/µl).

In tube 3 approximately 40 nanograms (4 µl) of CIP-treated plasmid was added to 16 µl of sterile distilled water.

Tubes 2 and 3 were control tubes to confirm the efficiency of CIP treatment and the presence of uncut plasmid, if any, respectively. The reaction mixtures were incubated overnight at 14°C and stored at -20°C for future use.

Preparation of Competent Cells

Sixty milliliters of LB media was inoculated with a single colony of *E. coli* HMS174 cells. Cells were grown exponentially to an OD_{590} of 0.5. At this point, the cells were spun at 5000 rpm (3,500 x g) for 10 minutes. The supernatant was discarded and the cells were resuspended in 10 ml of cold 0.1 M calcium chloride. The cells were incubated on ice for 10 minutes, followed by centrifugation at 5000 rpm (3,500 x g) for 10 minutes. The supernatant was discarded and the cells were resuspended in 2 ml of calcium chloride. At this point the cells were ready for transformation (Maniatis et al., 1982).

Transformation

Ten microliters of the ligation reaction was gently mixed with 0.2 ml of *E. coli* HMS174 cells. The tubes were incubated for 30 minutes on ice followed by heat-shock at 42°C for 2 minutes. The tubes were transferred on ice immediately and incubated for another 2 minutes. Eight hundred microliters of SOC media (95 ml distilled deionized water, 2 gm of bacto-tryptone, 0.5 gm of bacto-yeast extract, 50 gm NaCl, 1 ml of 250 mM KCl, pH adjusted to 7.0 with 1N NaOH. To the above sterilized media were added 0.5 ml of 2 mM $MgCl_2$ and 2 ml of sterile 1 M glucose) was added to each tube and incubated at 37°C for 45 minutes with gentle shaking. The tubes were spun for 5 minutes at 5000 rpm (3,500 x g). Eight hundred microliters of the supernatant was discarded and the cells were resuspended in the remaining media. Recombinant plasmids were selected on LB plates containing 50 µg/ml of ampicillin. The plates were incubated at 37°C overnight. The next day 5 recombinant plasmids were picked and each was inoculated into 5 ml of LB media containing 50 µg/ml of ampicillin. The cells were grown to saturation overnight and the next day plasmid DNA was isolated as described earlier in this chapter.

Restriction analyses was carried out to ensure the correct patterns. For the Glu287 mutants SpeI restriction digests to screen for positive mutant clones were carried out as follows: 2 µl plasmid DNA (approximately 20 nanograms) was added to 1 µl NEB buffer 1 (10 mM Bis Tris Propane-HCl, 10 mM magnesium chloride, 1 mM DTT, pH 7.0), 1 µl SpeI (1000 units/µl) and 6.5 µl sterile water. The tubes were incubated at 37°C for 2 hours and subjected to electrophoresis on a 1% high melt agarose gel. The positive clones were selected for DNA sequencing.

Double Stranded DNA Sequencing

DNA was sequenced by the Sanger dideoxy method (Sanger et al., 1977). Sequencing was carried out using standard protocols supplied by the United States Biochemical Corporation (USBC). Sequencing primers were synthesized close to the mutagenesis sites.

The sequence of the primers used for sequencing the Glu 13 mutants and the Glu287 mutants are listed below in Table 2-4.

Table 2-4. Primers synthesized for sequencing the Glu13 and Glu287 mutations.

Strand	Sequence (5' to 3')	Position
-	GTC TCT CCG TCC ATG GTG ATG CTG T	5344-5369
-	ACA GGT TGC TCG AGC CGG TGT CAA AGA	4845-4873
-	AGT ATC CAG GTA GTT CTC AAG	4800-4821
+	TAC TGG CAG ATT ACC TTG GAC	5350-5371
-	TGT TGG CCC TGT CAA AGA CGG	5717-5738

Approximately 5 µg of plasmid DNA (10 µl) was denatured using 10 µl sterile distilled water and 2 µl of freshly made 2M NaOH/2 mM EDTA. The tubes were incubated at 37°C for 30 minutes. The reaction was neutralized by the addition of 2 µl of 2M ammonium acetate pH 4.6. Denatured DNA was precipitated with 75 µl of cold absolute ethanol. The tubes were kept at -70°C for 2 hours, followed by centrifugation at 12,000 rpm (11,000 x g) for 10 minutes. The supernatant was discarded and the DNA pellet was rinsed with 100 µl of 75% cold ethanol. Centrifugation was repeated, supernatants discarded carefully and the tubes were dried in the 37°C incubator to get rid of the residual ethanol.

The Sequenase Version 2.0 Kit (USBC) was used for sequencing. All the buffers and solutions were supplied in the kit. To the dried DNA pellet was added 7 µl of sterile distilled water and 2 µl of Sequenase reaction buffer (200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl) and 1 µl of the desired primer (0.5 pmoles/µl). Annealing was allowed to proceed at 65°C for 2 minutes and subsequent cooling to 23°C over a 30 minute time period. The samples were now ready for labeling. The labeling reaction was started by making a master mix with 0.3 pmol/µl $[\alpha\text{-}^{35}S]$-ATP (Amersham) (1000 Ci/mmol), 20 mM DTT, 1.5 mM dGTP, 1.5 mM dCTP, 1.5 mM dTTP, 2 units of sequenase (13 units/µl) and incubating the tubes at room temperature for 5 minutes. At the end of 5 minutes, for each of the primer to be sequenced, the mixture containing the plasmid, primer and the sequenase enzyme, was divided into four tubes labeled "G", "A", "T" and "C". Each of these tubes contained 2.5 µl of the respective termination mixtures. The composition of the "C" termination mixture was as follows: 80 mM dGTP, 80 mM dATP, 80 mM dTTP, 80 mM dCTP, 8 mM ddCTP and 50 mM NaCl. The other termination mixtures had similar compositions. The tubes were incubated at 37°C for 15 minutes. This was followed by the TdT treatment (terminal deoxynucleotide transferase) (Kho and Zarbl, 1992).

To each of the tubes was added 1 µl of TdT mix (20.3 µl water, 6 µl TdT reaction buffer (200 mM Tris-HCl, 100 mM $MgCl_2$, 250 mM NaCl), 2.3 µl TdT (5 units/µl) and 1.5 µl of 20 mM dNTPs. The tubes were incubated at 37°C for 30 minutes. The reaction was terminated by the addition of 4 µl of stop solution (95% v/v formamide, 20 mM EDTA, 0.05% w/v Bromophenol blue, and 0.05% w/v xylene cyanol). The samples were stored at -70°C.

The reactions were heated at 100°C for 10 minutes prior to loading on a sequencing gel. Electrophoresis was carried out on a wedged gel consisting of 6% (w/v) polyacrylamide, 8 M urea, and 1 M TBE (10.8 gm of Tris base, 5.5 gm of boric acid, 0.93 gm of Na_2 EDTA, water to 100 ml), pH 8.3. The buffer used for electrophoresis was 1M TBE. Upon completion of electrophoresis the gel was fixed for 30 minutes in a solution containing 15% (v/v) glacial acetic acid, 15% (v/v) absolute ethanol and 70% water. The gel was then transferred to a 3M paper and covered with a plastic wrap. The gel was dried at 80°C for 1 hour on a gel drying apparatus (Bio-Rad). The gel was finally subjected to autoradiography on a XAR film for 48 hours at room temperature. The DNA sequence was manually read and the codon changes for mutagenesis were confirmed.

Expression of Glu13 and Glu287 Mutants

In order to express the Glu13 and Glu287 mutants, the pET-3b plasmid containing the mutant pepsinogen gene was transformed to an expression cell line *E. coli* BL21(DE3). Competent cells were made and transformation carried out as explained earlier in the chapter.

Cells containing the mutant plasmid were selected by plating the transformed cells on LB plates containing 50 μg/ml ampicillin. A single colony from these plates was used to inoculate 100 ml of M9CA media (787 ml of distilled deionized water, 100 ml of sterile 5% casamino acids, 100 ml of 10X M9 salts, (250 ml of distilled deionized water, 64 gm of Na_2HPO_4 7H2O, 15 gm of K_2HPO_4, 2.5 gm of NaCl and NH_4Cl), 2 ml of 1 M $MgSO_4$, 1 ml of 0.1 M $CaCl_2$ and 29.6 µl 1 M thiamine. Before use, glucose was added to the media at a 0.2% final concentration). The culture was grown to saturation overnight.

The next day, a 1:50 dilution was made into 4 liters of LB media containing 50 µg/ml ampicillin. Cells were grown at 37°C, to an $O.D_{600}$ of 0.5. At this time, the cells were induced with IPTG (final concentration 0.5 mM). The O.D of the cells were monitored at 1 and 3 hour after induction and 1 ml of the culture was aliquoted, spun at 5000 rpm (3,500 x g) and the pellet reconstituted in 100 µl of water and 100 µl of LSB (1.25 ml of 0.5 M Tris-HCl, pH 6.8, 1 ml 20% SDS, 500 µl β-mercaptoethanol, 1 ml glycerol and 250 µl 0.4% bromophenylblue). These samples were stored at 4°C to determine the relative levels of protein expression after induction using SDS-PAGE. At the end of 3 hours the cells were pelleted in pre-weighed bottles, for 10 minutes, at 5000 rpm (3,500 x g). The bottles were weighed and the total mass of the cells was determined. The bottles were stored in -20°C for further experiments.

SDS-PAGE Analysis of Expression of Protein

Samples were analyzed on a 10% Tris-Tricine SDS-polyacrylamide gels. Procedures for casting of gels was according to the standard Bio-Rad protocols. After electrophoresis the gels were fixed, stained and destained using the following modified procedures (Schagger and von Jagow, 1987).

Fixation. This step was very important since pepsin can be soluble in acids. The gel was fixed for 30 minutes in a mixture containing 40% absolute ethanol, 20% glacial acetic acid and 40% water.

Staining. The gel was stained for 1 hour in a mixture of 0.15% Coommassie R-250 (Bio-Rad), 30% ethanol, 10% glacial acetic acid, 60% water and 0.06% $CuSO_4$.

Destaining. The gel was destained in a solution containing 15% ethanol, 5% glacial acetic acid and 80% water, until the background of the gel was clear. The expression of the recombinant protein was confirmed,

Isolation of Inclusion bodies

To the pelleted cells was added 4.2 ml of TN buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl and 1 mM $MgCl_2$) per gram of cells. This was followed by the addition of 80 Kunits of DNAse (stock 60 Kunits/ml) per ml of cell suspension.

The suspension was incubated on ice for 10 minutes and subjected to two passes through a French press to lyse the cells. The lysed cells were carefully overlaid on a 27% sucrose cushion (10 ml) in a 30 ml corex tube and spun at 8500 rpm (11,300 x g) in a swinging bucket rotor for 30 minutes. The pellets were resuspended in 10 ml of TN buffer containing 1% Triton X-100. The suspension was overlaid on a 10 ml 27% sucrose cushion in a pre-weighed 30 ml corex tube and the centrifugation repeated. The supernatant was aspirated and the tubes were weighed to determine the wet weight of the inclusion bodies. The resulting pellet was stored at -20°C.

Refolding

Refolding of the wild-type and mutant recombinant pepsinogen in order to recover enzymatic activity involved denaturation, reduction and dialysis. The purified inclusion bodies were dissolved in freshly deionized 8 M urea, 50 mM CAPS pH 10.5, 1 mM EDTA, 1 mM glycine, 500 mM NaCl and 300 mM β-mercaptoethanol to a final concentration of approximately 1 mg/ml. After stirring at room temperature for 1 hour, the solution was centrifuged at 24,000 x g to remove precipitates. The supernatant was dialyzed in a SpectraPor 1 (MWCO 6-8 kDa) membrane against five times the original volume with a 50 mM Tris-HCl pH 11.0 buffer for 1 hour at room temperature. Following a buffer change and dialysis at room temperature for another hour, the dialysis buffer was changed to 50 mM Tris-HCl pH 7.5 and dialyzed overnight at

4°C. The next morning the buffer was changed to 50 mM MOPS pH 7.0 and dialyzed for at least 6 more hours at 4°C. Dialysis was repeated with 50 mM MOPS buffer to completely remove the β-mercaptoethanol. The resulting solution was centrifuged at 24,000 x g to remove precipitates and concentrated using a Millipore Ultrafiltration system fitted with low protein binding PLTK 10,000 NMWCO (molecular weight cutoff) membrane plates.

Gel Filtration Chromatography

The zymogen was further processed by centrifugation at 45,000 x g to remove precipitates before loading the protein solution onto a 2.5 cm x 90 cm Sephacryl S300 gel filtration column equilibrated with 20 mM MOPS pH 7.0 containing 300 mM NaCl. The zymogen was eluted at a flow rate of 0.4 ml/minute and the fractions showing the highest purity and activity were concentrated and buffer exchanged with ice cold water using an Amicon Centriprep10 unit. The purified zymogen was stored in 1 ml aliquots at -20°C.

Activation to Mature Form

In the activation of the native, Glu13Ala and Glu13Gln mutant pepsins, citric acid, pH 2.0, was added to the sample to give a final concentration 0.1 M, then held for 3 hours at 37°C. The Glu287Met, Glu287Gln and Glu13Gln/Glu287Met double mutants were activated with 0.1 M sodium formate buffer, 37°C, pH 3.5, for 24 hours.

Activation of pepsinogen for SDS-PAGE analysis was performed as above. Aliquots were taken and the reaction stopped at different time intervals by the addition of Laemmli sample buffer without glycerol and Coommassie stain (60 mM Tris-HCl, pH 6.8, 50 mM β-mercaptoethanol). The samples were then acetone precipitated and the proteins were resolved on 10% Tris-Tricine SDS-PAGE gels and stained with Coomassie Blue as described earlier in this chapter.

Aliquots of active enzyme were stored at -20°C until purification on an ion-exchange column.

Ion-Exchange Chromatography

The activated fractions were directly injected onto a Pharmacia Mono Q (anion exchange) column equilibrated with 20 mM MES pH 6.0. The enzyme was eluted by running a 35 minute gradient to 50% 20 mM MES pH 6.0 with 2M NaCl at 1 ml/min flow rate. Fractions containing activity were determined by the assay of cleavage of a good substrate Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu. Fractions with the highest activity were pooled and stored in 100 µl aliquots at -20°C.

N-Terminal Sequencing

The N-terminal sequence of the active, purified pepsin was confirmed using Applied Biosystems 470A and 473A Protein

Sequencers. The protein sample was applied to a ProSpin™ Cartridge which uses a ProBlott™ PVDF membrane attached to an insert in contact with a 3K Dalton cutoff filter. This enables the concentration of the protein with the removal of buffers, salts or detergents. The cartridge was spun at 5600 x g for 4-6 hours and PVDF membrane containing the protein was removed and submitted to the Protein Chemistry Core Facility for N-terminal sequencing.

Spectrophotometric Analyses

Substrate cleavage. Stock peptide solutions were made in distilled water and quantified by amino acid analysis (Beckman System 6300 High Performance Amino Acid Analyzer). The enzyme-catalyzed hydrolyses of the substrates were monitored at 37°C, pH 3.5, in a 0.1 M sodium formate buffer by averaging the decrease in absorbance over a range of 284-324 nm using a Hewlett Packard Diode Array Spectrophotometer. For all the kinetic assays, six tubes containing a mixture of the buffer and the enzyme were prewarmed for 3 minutes at 37°C. At the end of the preincubation, 6 different concentrations of the substrate were added and the absorbance was monitored using a multi-cell transporter (Scarborough and Dunn, 1994). The K_m and V_{max} values were determined from the initial rates using the equation.

$$v = \frac{V_{max} * [S]}{K_m + [S]}$$

With the same conditions as mentioned above, the cleavage of the peptides was also monitored by HPLC followed by sequence determination of the respective peaks to ensure the correct site of cleavage (Phe*Nph) by the enzyme. The pH dependence studies were carried out using the same procedure with formate/NaCl, acetate/NaCl or MES/NaCl ($I = 0.1$ M) at pH values ranging from 3 to 6.

Active site titration. The inhibitor stock solutions used for active site titrations were made in 100% DMSO and quantified by amino acid analysis. The amount of active enzyme used in the assay was determined by competitive inhibition of porcine pepsin by a potent Eli Lilly inhibitor LP-130 ($K_i = 0.5 \pm 0.05$ nM). The resulting curve was fitted with the Henderson equation on the Enzfitter program (Henderson, 1972).

The k_{cat} values (sec^{-1}) were determined by the equation

$$k_{cat} = \frac{V_{max}(AU/sec)}{\Delta A/[S]*[E]}$$

where ΔA = Total change in absorbance observed with 50 μM substrate [S]. [E] = total enzyme concentration (Molar). The errors were propagated by using standard procedures outlined by Meyer (1975).

pKa determination. The pKa values were calculated using plots of log k_{cat}/K_m vs pH in the Enzfitter program which uses the following equation to calculate pKa:

$$X = \frac{X_{(min)} + X_{(max)}.10^{(pH-pK_a)}}{1+10^{(pH-pK_a)}}$$

where $X = k_{cat}/K_m$, X(min) and X(max) are estimates of lower and upper limits of X

CHAPTER 3
SUBSTRATE SPECIFICITY REQUIREMENTS OF PORCINE PEPSIN: ANALYSES USING CHROMOGENIC OCTAPEPTIDE SUBSTRATES WITH SYSTEMATIC SUBSTITUTIONS IN THE P_5-P_3' POSITIONS

Introduction

The present chapter outlines the kinetic studies designed to explore potential interactions in the active site cleft of porcine pepsin using chromogenic octapeptide substrates with two bulky aromatic residues Phe and Nph in the P_1 and P_1' positions, respectively (Schechter and Berger, 1967, Fruton, 1970, Dunn, 1992). Eventhough all the members of the aspartic proteinase family of enzymes are structurally similar, they are functionally distinct as discussed in Chapter 1. These subtle distinctions arise due to the specific interactions between the solvent exposed amino acid residues that define the individual binding pockets and the side chains of the amino acids of the peptide that binds to the active site. These specific interactions confer the efficiency of proteolysis and the the subtle distinctions in the enzymatic specificity of all the members of this family. The substitutions of a series of amino acids in the P_5-P_3' (Table 3.1) positions were used to probe specific interactions such as hydrophobic attractions, van der Waals forces, hydrogen bonding and electrostatic interactions taking place in the individual binding pockets leading to catalysis for porcine pepsin

Also reported are the cation binding preferences of the S_3 and S_2 subsite of porcine pepsin, analyzed by enzyme-catalyzed hydrolysis of chromogenic octapeptide substrates at pH values ranging from 3 to 6. These studies are indicative of *specific* and/or *general* electrostatic interactions in the S_3 and S_2 subsite of porcine pepsin, involving the Glu13 and Glu287 residue, respectively. All the materials and methodology used in the kinetic characterization are discussed earlier, in detail, in Chapter 2.

Results

Kinetic Analysis

Table 3-1 lists the kinetic parameters of the substrates of the form Lys(P_5)-Pro(P_4)-Ala(P_3)-Lys(P_2)-Phe(P_1)*Nph(P_1')-Arg(P_2')-Leu(P_3') with systematic substitutions based on charge, size, hydrogen bonding and hydrophobicity at every position flanking the scissile bond (Phe*Nph).

All the peptides were cleaved by porcine pepsin at measurable rates at pH 3.5, with the notable exception of the peptide containing an arginine residue in the P_3 position.

The S_5 subsite accommodated all the substitutions tested. The lysine (0.8 $\mu M^{-1}s^{-1}$) and leucine (0.7 $\mu M^{-1}s^{-1}$)

residues exhibited a 3-fold preference in comparison with the hydrogen bonding serine (0.2 $\mu M^{-1}s^{-1}$)

The S_4 subsite had a preference for the proline residue in the control peptide (0.8 $\mu M^{-1}s^{-1}$) followed by the alanine (0.3 $\mu M^{-1}s^{-1}$). Powers et al. (1977) in a survey of cleavage junctions of pepsin specificity determined that proline was optimal only in the P_4 and P_3 positions.

This effect of proline substitution may be due to the restricted conformation of the peptide, thus aligning the peptide in a productive interaction.

Turning the focus to the S_3 subsite, the alanine in the control peptide (0.8 $\mu M^{-1}s^{-1}$) was the most productive in agreement with earlier reports (Dunn et al., 1986) with the highest k_{cat}/K_m values, followed by leucine (0.6 $\mu M^{-1}s^{-1}$). The serine (0.2 $\mu M^{-1}s^{-1}$) and aspartate (0.2 $\mu M^{-1}s^{-1}$) substitutions were 4-fold less acceptable than the control peptide.

The exceptional case was the poor cleavage of the substrate with an arginine residue in the P_3 position at the assay pH of 3.5. An inspection of the crystal structure of pepsin complexed with an inhibitor (A63218) (Abad-Zapatero et al., 1991) revealed that the prominent interactions taking place in the S_3 subsite of porcine pepsin are hydrophobic owing to the residues Tyr75, Thr77, Phe111, Ala115, Pro116, Gly217, and Thr218. The hydrophobic nature of the pocket can enhance hydrophobic interactions with alanine and leucine, leading to productive binding. Another residue within the S_3

Table 3-1 Kinetic parameters of peptide substrates of the form Lys-Pro-Ala-Lys-Phe*Nph[a]-Arg-Leu with substitutions in P_5 to P_3' positions

P_5	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	k_{cat} (sec^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Lys	Pro	Ala	Lys	Phe	Nph	Arg	Leu	65 ± 10	83 ± 8	0.80 ± 0.09
Ala	Pro	Ala	Lys	Phe	Nph	Arg	Leu	5 ± 1	15 ± 3	0.30 ± 0.08
Asp	Pro	Ala	Lys	Phe	Nph	Arg	Leu	7 ± 1	15 ± 2	0.50 ± 0.08
Ser	Pro	Ala	Lys	Phe	Nph	Arg	Leu	2 ± 0.3	9 ± 3	0.20 ± 0.08
Leu	Pro	Ala	Lys	Phe	Nph	Arg	Leu	6 ± 1	9 ± 1	0.70 ± 0.09
Arg	Pro	Ala	Lys	Phe	Nph	Arg	Leu	8 ± 1	33 ± 7	0.40 ± 0.06
Lys	*Ala*	Ala	Lys	Phe	Nph	Arg	Leu	50 ± 12	168 ± 30	0.30 ± 0.09
Lys	*Asp*	Ala	Lys	Phe	Nph	Arg	Leu	8 ± 1	208 ± 29	0.04 ± 0.01
Lys	*Ser*	Ala	Lys	Phe	Nph	Arg	Leu	11 ± 1	124 ± 17	0.09 ± 0.01
Lys	*Leu*	Ala	Lys	Phe	Nph	Arg	Leu	10 ± 1	159 ± 68	0.06 ± 0.03
Lys	*Arg*	Ala	Lys	Phe	Nph	Arg	Leu	9 ± 1	238 ± 20	0.04 ± 0.01
Lys	Pro	*Asp*	Lys	Phe	Nph	Arg	Leu	45 ± 6	270 ± 37	0.20 ± 0.03
Lys	Pro	*Ser*	Lys	Phe	Nph	Arg	Leu	10 ± 1	57 ± 10	0.20 ± 0.04
Lys	Pro	*Leu*	Lys	Phe	Nph	Arg	Leu	52 ± 6	81 ± 8	0.60 ± 0.10
Lys	Pro	*Arg*	Lys	Phe	Nph	Arg	Leu	PC[b]	PC	PC
Lys	Pro	Ala	*Ala*	Phe	Nph	Arg	Leu	31 ± 4	40 ± 7	0.80 ± 0.17
Lys	Pro	Ala	*Asp*	Phe	Nph	Arg	Leu	43 ± 7	84 ± 20	0.50 ± 0.07
Lys	Pro	Ala	*Ser*	Phe	Nph	Arg	Leu	30 ± 3	58 ± 6	0.50 ± 0.07
Lys	Pro	Ala	*Leu*	Phe	Nph	Arg	Leu	13 ± 3	28 ± 9	0.70 ± 0.20
Lys	Pro	Ala	*Arg*	Phe	Nph	Arg	Leu	17 ± 2	35 ± 6	0.50 ± 0.10
Lys	Pro	Ala	Lys	Phe	Nph	*Ala*	Leu	29 ± 3	48 ± 7	0.60 ± 0.11
Lys	Pro	Ala	Lys	Phe	Nph	*Asp*	Leu	5 ± 1	106 ± 36	0.06 ± 0.02
Lys	Pro	Ala	Lys	Phe	Nph	*Ser*	Leu	25 ± 3	48 ± 11	0.50 ± 0.10
Lys	Pro	Ala	Lys	Phe	Nph	*Leu*	Leu	23 ± 4	21 ± 7	1.0 ± 0.40
Lys	Pro	Ala	Lys	Phe	Nph	Arg	*Ala*	10 ± 2	80 ± 18	0.13 ± 0.03
Lys	Pro	Ala	Lys	Phe	Nph	Arg	*Asp*	12 ± 2	127 ± 19	0.09 ± 0.01
Lys	Pro	Ala	Lys	Phe	Nph	Arg	*Ser*	8 ± 1	99 ± 23	0.08 ± 0.02
Lys	Pro	Ala	Lys	Phe	Nph	Arg	*Arg*	7 ± 1	43 ± 7	0.15 ± 0.03

[a]Nph = p -nitrophenylalanine, [b]PC = Poorly cleaved under standard assay conditions (0.1 M sodium formate buffer, pH 3.5, 37°C)

subsite is Glu13. pH dependence studies of the arginine containing peptide revealed a 20-fold increase in the rate of cleavage with a titratable group having, on average, pK_a of 4.2. A candidate for this group was the Glu13 residue. At a pH of 3.5, the glutamate side chain has hydrophobic character ascribed to the methylene group in the side chain. However, at a higher pH, deprotonation would render the glutamic acid side chain negatively charged and it could participate in electrostatic interactions with the positively charged arginine side chain of the substrate. Detailed pH dependence studies are discussed later in this chapter.

Turning to the results of the P_2 substituents in Table 3-1, the energy contributed by the binding of all the substrates are comparable with each other. As indicated by the k_{cat}/K_M values, the enzyme subsite seems to accommodate hydrophobic, charged, as well as hydrogen bonding residues productively (0.5 - 0.8 $\mu M^{-1}s^{-1}$). Based on the crystal structure data, the S_2 subsite of porcine pepsin is predominantly hydrophobic. A residue which borders the S_2 and the S_4 subsite is Glu287 which could be involved in electrostatic interactions with a positively charged residue. This residue is variable among this family of enzymes, with Gln, Met and Gly occurring in various enzymes. The pH dependent changes in kinetic parameters, indicative of electrostatic interactions owing to the Glu287 residue, in the S_2 subsite, have been reported earlier (Pohl and Dunn, 1988).

The S_2' pocket accepts hydrophobic residues such as leucine optimally exhibiting the highest k_{cat}/K_m with all the substitutions tested (1.0 $\mu M^{-1}s^{-1}$). Arginine, serine and alanine exhibited comparable numbers (0.5 - 0.6 $\mu M^{-1}s^{-1}$). The least acceptable residue was the negatively charged aspartate (0.06 $\mu M^{-1}s^{-1}$).

Similar studies carried out with the same set of peptides with cathepsin E, another member of this family, presented comparable low tolerance of aspartate in the P_2' position. This suggests that the negative charge for both these enzymes resulted in either reduced binding or a 10-fold decrease in catalysis.

The final substitution under consideration was the P_3' position. The substitutions were not particularly discriminatory. There was only a 2-fold decrease in the k_{cat}/K_m values for the serine substitution in comparison to the leucine in the control peptide.

Examining all the substitutions, the major effect is seen with the charged substituents. This suggests that the overall charge of the binding cleft can influence the catalytic process.

Earlier kinetic studies on selective aspartic proteinases with the peptide Lys-Pro-X-Glu-Phe*Nph-Arg-Leu with substitutions in the P_3 position (Dunn et al., 1986) indicated that the most effective substrates for porcine pepsin had alanine and isoleucine replacements in the P_3 position. In order to understand the S_2 subsite preferences we designed chromogenic octapeptide substrates of the form Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu with systematic substitutions in the P_2 position shown in Table 3-2.

As evident from Table 3-2, the S_2 subsite accommodates hydrophobic substitutions readily which include residues such as leucine, alanine, norleucine, as well as β-branched isoleucine and valine (0.5 - 1.4 $\mu M^{-1}s^{-1}$).

The negatively charged aspartate also fits about as well with k_{cat}/K_m values approaching the control peptide (0.9 $\mu M^{-1}s^{-1}$). These observations can be correlated to the crystal structure solved by Abad-Zapatero et al. (1991). Based on the three-dimensional structure one can see a predominantly hydrophobic pocket in the S_2 subsite consisting of Thr77, Thr218, Met289 and Ile300. Another residue which borders the region between the S_2 and the S_4 subsite is Glu287, which potentially could be involved in electrostatic interaction.

The peptide containing histidine and lysine substitutions exhibited relatively low k_{cat}/K_m numbers. An interesting observation was the acceptance of arginine substitution (0.82 $\mu M^{-1}s^{-1}$) to a favorable extent as opposed to the poor acceptance of histidine (0.24 $\mu M^{-1}s^{-1}$) and lysine (0.2 $\mu M^{-1}s^{-1}$). pH dependence studies (Table 3-3) were carried out with both the peptides with arginine and lysine in the P_2 position to see whether any of these residues had any observable control over the speculated deprotonation of the Glu287 residue positioned between the S_2 and S_4 subsites.

Table 3-2. Kinetic parameters of peptide substrates of the form Lys-Pro-Ile-Glu-Phe*Nph^a-Arg-Leu with systematic substitutions in the P_2 position

P_5	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	k_{cat} (sec^{-1})	K_m (μM)	k_{cat}/K_m (μM^{-1}s^{-1})
Lys	Pro	Ile	*Leu*	Phe	Nph	Arg	Leu	20 ± 6	14 ± 5	1.40 ± 0.60
Lys	Pro	Ile	*Ala*	Phe	Nph	Arg	Leu	30 ± 9	25 ± 8	1.20 ± 0.50
Lys	Pro	Ile	*Nle*	Phe	Nph	Arg	Leu	15 ± 4	13 ± 3	1.14 ± 0.40
Lys	Pro	Ile	*Glu*	Phe	Nph	Arg	Leu	72 ± 9	70 ± 11	1.02 ± 0.20
Lys	Pro	Ile	*Ser*	Phe	Nph	Arg	Leu	57 ± 17	63 ± 18	0.91 ± 0.40
Lys	Pro	Ile	*Asp*	Phe	Nph	Arg	Leu	46 ± 12	52 ± 9	0.88 ± 0.30
Lys	Pro	Ile	Arg	Phe	Nph	Arg	Leu	20 ± 5	24 ± 4	0.82 ± 0.30
Lys	Pro	Ile	*Ile*	Phe	Nph	Arg	Leu	14 ± 4	17 ± 6	0.81 ± 0.50
Lys	Pro	Ile	*Val*	Phe	Nph	Arg	Leu	17 ± 5	31 ± 8	0.56 ± 0.20
Lys	Pro	Ile	*Asn*	Phe	Nph	Arg	Leu	22 ± 6	43 ± 4	0.51 ± 0.10
Lys	Pro	Ile	*Gln*	Phe	Nph	Arg	Leu	24 ± 8	72 ± 28	0.33 ± 0.10
Lys	Pro	Ile	*His*	Phe	Nph	Arg	Leu	8 ± 2	33 ± 5	0.24 ± 0.07
Lys	Pro	Ile	*Thr*	Phe	Nph	Arg	Leu	17 ± 5	80 ± 12	0.22 ± 0.06
Lys	Pro	Ile	*Lys*	Phe	Nph	Arg	Leu	5 ± 1	28 ± 7	0.18 ± 0.06
Lys	Pro	Ile	Pro	Phe	Nph	Arg	Leu	NC^b	NC^b	NC^b

^{a}Nph = p -nitrophenylalanine, ^{b}NC = Not cleaved under standard assay conditions (0.1M sodium formate, pH 3.5, 37°C).

Proline was totally unacceptable in the S_3 subsite which may be explained due to the unique conformation of the proline side chain leading to unproductive binding. This is in agreement with the studies on pepsin carried out by Powers et al. (1977)

pH Dependence Studies

To understand the specific cation binding preferences in the S_3 and S_2 subsites specifically, pH dependence studies were carried out with the following peptides. These peptides have a positively charged residue in either the P_2, P_3 or both the positions. Peptide 6 was used as a control peptide with no positive charges in both the P_2 and P_3 position of the peptide.

1. Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu.
2. Lys-Pro-Arg-Lys-Phe*Nph-Arg-Leu
3. Lys-Pro-Arg-Ala-Phe*Nph-Arg-Leu
4. Lys-Pro-Ile-Lys-Phe*Nph-Arg-Leu
5. Lys-Pro-Ile-Arg-Phe*Nph-Arg-Leu
6. Lys-Pro-Ile-Ala-Phe*Nph-Arg-Leu

The result is schematically represented in Figure 3-1 showing the titration curve of log k_{cat} as a function of pH for the peptide 2.

Table 3-3 represents the pK_a values obtained from graphs of k_{cat} or k_{cat}/K_m versus pH for the six peptides under study.

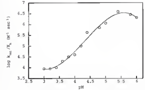

Figure 3-1. pH dependence curve of the peptide Lys-Pro-Arg-Lys-Phe*Nph-Arg-Leu with wild type porcine pepsin.

Table 3-3 pK_a values of enzyme-catalyzed cleavage of peptide substrates of the form Lys-Pro-Xaa-Yaa-Phe*Nph[a]-Arg-Leu

P_5	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	k_{cat} vs. pH	k_{cat}/K_m vs. pH
Lys	Pro	Ala	*Lys*	Phe	Nph	Arg	Leu	-----	4.7 ± 0.1
Lys	Pro	Arg	*Lys*	Phe	Nph	Arg	Leu	4.4 ± 0.2	5.3 ± 0.3
Lys	Pro	Arg	Ala	Phe	Nph	Arg	Leu	4.0 ± 0.2	5.0 ± 0.1
Lys	Pro	Ile	*Lys*	Phe	Nph	Arg	Leu	-----	4.9 ± 0.1
Lys	Pro	Ile	*Arg*	Phe	Nph	Arg	Leu	-----	5.4 ± 0.3
Lys	Pro	Ile	Ala	Phe	Nph	Arg	Leu	NT[b]	NT[b]

[a] Nph = *p*-nitrophenylalanine

[b] NT = No titratable response observed

It is evident from the k_{cat} vs. pH profile of the wild type enzyme with peptides 2 and 3 with an arginine in the P_3 position that the group that is titrated has a pK_a of around 4.2. Since this pK_a is associated with the peptides containing an arginine in the P_3 position we can attribute this pK_a to the Glu13 residue in the S_3 subsite of the enzyme.

The k_{cat}/K_m parameter for peptides 4 and 5 exhibits an apparent pK_a of around 5.2. As discussed earlier, the S_2 subsite is predominantly composed of hydrophobic amino acids with the exception of the more polar Glu287.

At this juncture, this apparent pK_a can be ascribed to the Glu287 residue which has the ability to provide electrostatic attraction for positively charged residues in the P_2 position of the substrate at a higher pH. No titratable response was observed for the control peptide with an alanine in the P_2 position and a isoleucine in the P_3 position.

Discussion

The S_3 subsite of porcine pepsin is defined by the residues Tyr75, Thr77, Phe111, Tyr114, Ala115, Pro116, Gly217, and Thr218 (Abad-Zapatero et al., 1991). The hydrophobic nature of the pocket can enhance hydrophobic interactions with the β-branched bulky isoleucine and the hydrophobic alanine residue to the same extent. The peptides may also anchor strongly owing to the size of the S_3 pocket Glu 13 in the S_3 pocket of the enzyme may be involved in specific electrostatic interactions with a positive charge in P_3 position of the substrate.

Position 287 can greatly influence the binding of the substrates with a positive charge in the P_2 position for porcine pepsin. Glu287 in porcine pepsin which borders the S_2 and S_4 subsite can be involved in electrostatic interactions with the positively charged lysine residue.

Mutations at Glu13 and Glu287 positions were carried to test the hypothesis. Detailed kinetic characterization of the mutant enzymes are discussed in Chapter 4.

In conclusion, the replacements in the P_3 and P_2 positions significantly affect catalysis. Of importance is the presence of two glutamic acid residues unique to porcine pepsin promoting electrostatic interactions in the S_2 and the S_3 subsites (refer to sequence alignment in Table 1-1). The evaluation of the role of these two residues is important in understanding the contribution of electrostatic interactions in enzymatic specificity. The subsequent chapters will discuss the mutagenesis results, where the Glu13 and Glu287 residues were changed to residues existing in either cathepsin D or cathepsin E to assess the contribution of electrostatic interactions in catalysis.

CHAPTER 4
ROLE OF GLU13 AND GLU287 RESIDUES IN GENERAL/SPECIFIC ELECTROSTATIC INTERACTIONS

Introduction

Based on the sequence alignment and comparing the three-dimensional crystal structures of the mammalian enzymes porcine pepsin, human cathepsin D, and human cathepsin E, two unique glutamic acid residues were present in porcine pepsin in the S_2 and S_3 subsites. These were Glu13 in the S_3 subsite and Glu287 in the S_2 subsite. In order to investigate the role of these two residues in the cation binding preferences and catalysis, mutants were generated at these positions replacing the residues with amino acids existing in human cathepsin D, human cathepsin D or the fungal enzyme endothiapepsin. An alanine (present in endothiapepsin) or a glutamine (present in cathepsin D) were introduced at position 13. A methionine (in cathepsin D) or a glutamine (in cathepsin E) were introduced at position 287.

In addition to the single mutations, a double mutant was also created with a glutamine at position 13 and a methionine at position 287 to mimic some of the enzymatic specificity of cathepsin D.

The methodology used to create the mutations and kinetic characterizations are detailed in Chapter 2. This chapter addresses the results of expression, activation and kinetic characterization of the mutants.

Results

Expression of Porcine Pepsinogen

Expression of the protein in *E. coli* BL21-DE3 cells was monitored using SDS-PAGE analysis as shown in Figure 4-1. Increased levels of expression was observed at 1 hour after induction and steadily increased up to 3 hours until the cells were pelleted. At every time point, the same number of cells were loaded on the SDS-PAGE gel to determine the relative protein levels expressed after induction.

Activation

Activation conditions for each of the mutants had to be optimized. Changes in time, pH and temperature were carried out independently in order to activate the pepsinogen to mature pepsin with the correct N-terminus (Ile-Gly-Asp -).

Activation of the Glu13 mutants was carried out at 37°C, pH 2.0, 0.1 M citrate buffer, for 3 hours. The Glu287 mutants and the double mutant were activated at 37°C, pH 3.5, 0.1 M formate buffer, for 24 hours. The activation profile of the Glu287Gln mutant is shown in Figure 4-2.

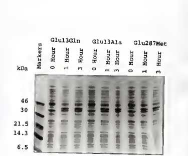

Figure 4-1. SDS-PAGE analysis of the expression of recombinant mutant Glu13Ala, Glu13Gln and Glu287Met porcine pepsinogen. Increased levels of expression of the mutant proteins are observed at one and three hours after induction (44 kDa).

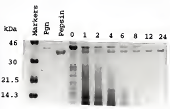

Figure 4-2 SDS-PAGE analysis of the activation of Glu287Gln mutant pepsinogen at pH 3.5, 37°C.

Purification

The samples prior to preparative activation were subjected to gel filtration chromatography to obtain purified pepsinogen. The fractions containing activity were determined by the assay of cleavage of a good substrate, Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu. Under the conditions of the assay, zymogen is instantaneously converted to active enzyme. The fractions with the highest activity were pooled and stored at -20°C to be activated. The activity elution profile of Glu13Ala pepsinogen is shown in Figure 4-3.

After activation, the samples were further purified using anion exchange chromatography. A typical elution profile is shown in Figure 4-4. Fractions with the highest activity were pooled and stored at -20°C for kinetic characterization. Purity was assessed by mass spectrometry and SDS-PAGE analyses.

N-Terminal Sequencing

The activated enzyme was subjected to N-terminal sequencing, described in Chapter 2, to ensure that the enzyme had activated to mature form. All the mutants and the wild type pepsinogen activated to a mixture of two populations one with the mature sequence Ile-Gly-Asp--(80%) and the second with the longer sequence Ala-Ala-Leu-Ile-Gly-Asp-- (20%).

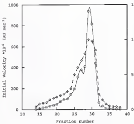

Figure 4-3. Gel filtration chromatography elution profile of Glu13Ala mutant enzyme

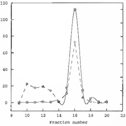

Figure 4-4. Anion exchange chromatography elution profile of Glu13Ala mutant enzyme

The second population was 3 amino acids longer than the mature form (Ile-Gly-Asp -). This did not have any effect on the kinetic parameters. Data to support this conclusion are presented in the Table 4-1 where the native and the wild type recombinant enzymes are compared. The table lists peptides of the form Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu tested with the wild type and wild-type recombinant enzyme.

Minor sequences were also seen for the Glu287 mutants. These were analyzed on a Perseptive Biosystem MALDI TOF (Matrix-Assisted Laser Desorption Ionization Time of Flight) Mass spectrometer, at the Protein Chemistry Core, at the University of Florida. The predominant peak was 35,000 daltons (90%) which is mature pepsin and the two minor peaks were 6000 and 4000 daltons, respectively.

Kinetic Analyses

Glu13 Mutations. In order to examine the changes in specificity, if any, by replacing an alanine or a glutamine at position 13, kinetic parameters (k_{cat}, K_m, k_{cat}/K_m) for the substrates of the form Lys-Pro-Ala Lys-Phe*Nph-Arg-Leu with substitutions in the P_3 and P_2 positions were determined. These results in comparison to the wild type enzyme are listed in Table 4.2. k_{cat}/K_m is the second order rate constant that defines all the steps of an enzyme reaction going from the ground state to the first irreversible step in the transition-state complex.

Table 4-1. Kinetic parameters of peptide substrates of the form Lys-Pro-Xaa-Yaa-Phe-Nph[a]-Arg-Leu with substitutions in P_3 and P_2 for wild type and wild type recombinant porcine papain

P_5	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	Wild Type Enzyme k_{cat} (sec^{-1})	Wild Type Enzyme K_m (μM)	Wild Type Enzyme k_{cat}/K_m (μM^{-1}s^{-1})	Wild Type recombinant enzyme k_{cat} (sec^{-1})	Wild Type recombinant enzyme K_m (μM)	Wild Type recombinant enzyme k_{cat}/K_m (μM^{-1}s^{-1})
Lys	Pro	Ala	Lys	Phe	Nph	Arg	Leu	65 ± 15	83 ± 8	0.8 ± 0.1	63 ± 7	90 ± 10	0.7 ± 0.07
Lys	Pro	Leu	Lys	Phe	Nph	Arg	Leu	52 ± 6	81 ± 8	0.6 ± 0.1	45 ± 8	101 ± 16	0.4 ± 0.08
Lys	Pro	Ser	Lys	Phe	Nph	Arg	Leu	10 ± 1	57 ± 10	0.2 ± 0.04	16 ± 3	52 ± 9	0.3 ± 0.08
Lys	Pro	Asp	Lys	Phe	Nph	Arg	Leu	45 ± 8	270 ± 37	0.2 ± 0.03	33 ± 6	230 ± 14	0.1 ± 0.02
Lys	Pro	Arg	Lys	Phe	Nph	Arg	Leu	PC[b]	PC[b]	<0.02	PC[b]	PC[b]	<0.02
Lys	Pro	Ala	Asp	Phe	Nph	Arg	Leu	45 ± 7	84 ± 20	0.5 ± 0.1	38 ± 7	110 ± 14	0.3 ± 0.05
Lys	Pro	Ala	Ala	Phe	Nph	Arg	Leu	31 ± 4	40 ± 7	0.8 ± 0.1	28 ± 3	34 ± 8	0.8 ± 0.10
Lys	Pro	Ala	Ser	Phe	Nph	Arg	Leu	30 ± 3	56 ± 6	0.5 ± 0.07	25 ± 5	79 ± 14	0.3 ± 0.06
Lys	Pro	Ala	Leu	Phe	Nph	Arg	Leu	19 ± 3	28 ± 3	0.7 ± 0.2	22 ± 8	25 ± 3	0.9 ± 0.20
Lys	Pro	Ala	Arg	Phe	Nph	Arg	Leu	17 ± 2	35 ± 8	0.5 ± 0.1	11 ± 2	30 ± 2	0.3 ± 0.05

[a] Nph = p-nitrophenylalanine
[b] PC = Poorly cleaved at pH 3.5

Table 4-2 Analysis of Glu13, Glu287 and double mutants of porcine pepsin at pH 3.5, with the substrate K(Lys)-P(Pro)-A(Ala)-K(Lys)-F(Phe)*X(Nph)^a-R(Arg)-L(Leu) with systematic substitutions in the P$_2$ and P$_3$ positions

PEPTIDE	WILD TYPE ENZYME			GLU13ALA MUTANT ENZYME			GLU13GLN MUTANT ENZYME		
	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
K P A K F*X R L	61 ± 10	63 ± 8	0.6 ± 0.1	57 ± 12	104 ± 20	0.5 ± 0.1	41 ± 7	112 ± 17	0.4 ± 0.08
K P R K F*X R L	—	—	<0.02	---	---	<0.02	—	—	<0.02
K P D K F*X R L	45 ± 6	270 ± 37	0.2 ± 0.03	12 ± 2	20 ± 3	0.6 ± 0.1	9 ± 1	40 ± 6	0.2 ± 0.05
K P E K F*X R L	18 ± 1	57 ± 10	0.3 ± 0.08	16 ± 3	72 ± 17	0.2 ± 0.07	40 ± 7	154 ± 10	0.3 ± 0.06
K P L K F*X R L	52 ± 6	61 ± 8	0.6 ± 0.1	13 ± 2	20 ± 2	0.6 ± 0.1	34 ± 5	58 ± 3	0.3 ± 0.06
K P A A F*X R L	31 ± 4	40 ± 7	0.9 ± 0.1	71 ± 13	38 ± 5	2.0 ± 0.4	60 ± 9	70 ± 17	0.7 ± 0.02
K P A R F*X R L	17 ± 2	35 ± 6	0.5 ± 0.1	22 ± 4	57 ± 11	0.4 ± 0.1	40 ± 5	71 ± 9	0.6 ± 0.10
K P A D F*X R L	43 ± 7	84 ± 20	0.5 ± 0.1	78 ± 16	132 ± 20	0.6 ± 0.1	21 ± 3	64 ± 8	0.3 ± 0.05
K P A E F*X R L	30 ± 3	58 ± 6	0.5 ± 0.07	90 ± 19	114 ± 20	0.8 ± 0.2	13 ± 2	106 ± 14	0.1 ± 0.02
K P A L F*X R L	19 ± 3	26 ± 9	0.7 ± 0.2	38 ± 7	27 ± 5	1.4 ± 0.3	35 ± 5	91 ± 17	0.3 ± 0.09

^aNph = p-nitrophenylalanine

Table 4-2 continued

PEPTIDE	GLU287MET MUTANT			GLU287GLN MUTANT			GLU12GLN/GLU287MET DOUBLE MUTANT		
	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
K P A K F*X E L	62 ± 8	95 ± 19	0.7 ± 0.1	80 ± 14	86 ± 10	0.9 ± 0.1	77 ± 14	250 ± 54	0.1 ± 0.08
K P E K F*X E L	---	---	<0.02	---	---	<0.01	---	---	<0.01
K P E K F*X R L	31 ± 8	307 ± 110	0.1 ± 0.04	59 ± 9	182 ± 34	0.3 ± 0.08	17 ± 5	303 ± 110	0.1 ± 0.01
K P E K F*E E L	47 ± 6	195 ± 35	0.2 ± 0.05	72 ± 11	144 ± 16	0.5 ± 0.1	10 ± 4	241 ± 22	0.1 ± 0.01
K P L E F*E E L	48 ± 5	117 ± 15	0.4 ± 0.06	89 ± 17	124 ± 17	0.7 ± 0.1	37 ± 4	211 ± 24	0.2 ± 0.03
K P A A F*X R L	171 ± 18	40 ± 4	4.0 ± 0.5	303 ± 54	61 ± 9	5.0 ± 1.0	85 ± 8	84 ± 14	1.0 ± 0.1
E P A E F*E R L	114 ± 11	52 ± 6	2.0 ± 0.3	190 ± 35	47 ± 6	4.0 ± 0.8	33 ± 4	106 ± 11	0.1 ± 0.04
K P A E F*X E L	153 ± 26	63 ± 10	3.0 ± 0.6	225 ± 38	43 ± 4	5.0 ± 0.8	63 ± 8	101 ± 12	0.7 ± 0.1
E E A S F*X E L	169 ± 19	78 ± 4	2.0 ± 0.3	365 ± 42	118 ± 10	1.0 ± 0.5	67 ± 7	137 ± 11	0.5 ± 0.06
K P A L F*X E L	90 ± 11	48 ± 8	2.0 ± 0.4	125 ± 17	40 ± 8	1.0 ± 0.8	11 ± 5	95 ± 20	0.1 ± 0.1

X = p-nitrophenylalanine

It is a universal measure to understand the specificity of an enzyme-substrate pair.

In considering the P_3 substitutions, the specificity constants (k_{cat}/K_m) did not exhibit any significant change in either mutant with respect to the wild type enzyme (average value 0.4 $\mu M^{-1}sec^{-1}$). The only exception was the substitution of aspartate in the P_3 position of the peptide which resulted in a three-fold increase in k_{cat}/K_m for the Glu13Ala mutant. The effect was mainly due to a 10-fold decrease in the Michaelis constant (K_m).

A clearly measurable improvement in the specificity constant for the Glu13Ala mutant was observed for the peptide substrates with the alanine (2 $\mu M^{-1}sec^{-1}$) and leucine (1.4 $\mu M^{-1}sec^{-1}$) substitutions in the P_2 position enzyme. No significant changes were observed for the Glu13Gln mutant. The other substitutions in the P_2 positions were not particularly discriminatory.

The peptide with an arginine at the P_3 position was not cleaved at pH 3.5 by either of the mutants. As noted in Chapter 3, the wild type enzyme also demonstrated poor cleavage of the P_3 arginine peptide at pH 3.5. Increasing the pH of the assay to 5.0 did not improve the catalytic efficiency of either the Glu13Ala or Glu13Gln mutant toward the arginine containing peptide. These results were in contrast to the wild type enzyme which exhibited a 20-fold increase in rates at higher pH, as reported in Chapter 3.

Glu287 and Double Mutants. Kinetic characterization of the Glu287 mutants and the double mutant (Glu13Gln/Glu287Met) is also outlined in Table 4-2. All the peptides were cleaved efficiently by all the mutants constructed. All the P_2 substituents were cleaved by the Glu287 mutants 3- to 6-fold more efficiently than by the wild type enzyme (Table 4-2). This increase in specificity constants was due to increase (approximately 3- to 6-fold) in the turnover number (k_{cat}) and not so much with the Michaelis constant (K_m). The double mutant exhibited specificity constants similar to the wild type enzyme.

pH Dependence Studies. To understand the cation binding preferences of the mutant enzymes in the S_3 and the S_2 subsite, pH dependence studies were carried out with the following peptides. These peptides were chosen because they have a positively charged residue in either the P_2 or the P_3 position (peptide 1 and peptide 2). The studies will aid in our understanding of the proposed electrostatic interactions taking place in the S_2 and S_3 subsites respectively. The results are presented in Table 4-3.

1. Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu
2. Lys-Pro-Arg-Ala-Phe*Nph-Arg-Leu
3. Lys-Pro-Ile-Ala-Phe*Nph-Arg-Leu

Table 4-3 pK$_a$ values of enzyme-catalyzed cleavage of peptide substrates of the form Lys-Pro-Xaa-Yaa-Phe*Nph^a-Arg-Leu

	Glu13Ala	Glu13Gln	Glu287Met	Glu287Gln	Glu13Gln/ Glu287Met
PEPTIDE	k_{cat}/K_m vs. pH	k_{cat}/K_m vs. pH	k_{cat}/K_m vs. pH	k_{cat}/K_m vs. pH	k_{cat}/K_m vs. pH
Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu	4.3	3.9	4.1	4.5	3.9
Lys-Pro-Arg-Ala-Phe*Nph-Arg-Leu	NT[b]	NT	4.1	4.3	ND[c]
Lys-Pro-Ile-Ala-Phe*Nph-Arg-Leu	NT	NT	NT	NT	NT

[a] Nph = p-nitrophenylalanine
[b] NT = No titratable response observed
[c] ND = Not determined

The Glu13 mutants did not exhibit any titratable response with the arginine containing peptide in the P_3 position as discussed earlier. The P_2 lysine containing peptide (peptide 1) exhibited an apparent pK_a around of 4.2, also with the Glu13 mutants. As discussed earlier, in Chapter 3, this pK_a can be attributed to the Glu287 residue which can permit electrostatic interactions at a higher pH.

However, the k_{cat}/K_m parameter also exhibits a pK_a of around 4.2 for all the Glu287 mutants tested with the peptide containing a lysine in the P_2 position (peptide 1). Since the Glu287 residue has been replaced by a non-titratable group, the observed pK_a can be attributed to the general long range electrostatic effects in the active site of porcine pepsin. These effects will be discussed in detail later in the chapter. The titratable response was due to improvement in Michaelis constant (K_m) as well as a small increase in the turnover number (k_{cat}).

Due to very high K_m values (>500 μM) it was difficult to test the pH dependent changes in kinetic parameters for the double mutant with the arginine containing peptide. The k_{cat}/K_m versus pH profile for the peptide containing a lysine in the P_2 position (peptide 1) exhibited a pK_a of around 3.9.

Peptide 3 showed no titratable response with any mutant, in agreement with earlier results for the wild-type enzyme.

K_i determinations. A competitive inhibitor was used to provide information on the selectivity of the mutant enzymes with respect to the wild type enzyme. Table 4-4 lists the K_i values determined for the Eli Lilly inhibitor LP-130 used in active site titrations for all the enzymes.

As evident from the data, the inhibitor binds more strongly (16-fold) to the Glu287 mutants in comparison to the wild type enzyme. This can be attributed to the stronger hydrophobic S_2 pocket in the mutant enzyme. A hydrophobic residue in the P_2 position of the inhibitor can anchor strongly to the S_2 subsite due to enhanced hydrophobic interactions. Similar hydrophobic effects are evident with the Glu13 mutant. The Glu13Gln mutant also has a 5-fold preference in the K_i values compared to the wild type enzyme.

Table 4-4 K_i values with a competitive inhibitor LP-130

ENZYME	K_i (nM)
Wild type	0.5
Glu13Gln	0.1
Glu287Met	0.03
Glu287Gln	0.03
Glu13Gln/Glu287Met	0.1

Discussion

The S3 Subsite

Replacing the Glu13 residue, either with an alanine or a glutamine, did not exhibit remarkable changes in enzymatic specificity. The k_{cat}/K_m values are similar within standard errors. However, this change was detrimental for the arginine containing peptide. Increasing the pH did not improve the catalytic efficiency of the enzyme toward the peptide. This was primarily due to the loss of the electrostatic interaction provided by the Glu13 residue in the wild type enzyme. This data is supportive of the hypothesis that the Glu13 residue is involved in specific electrostatic interactions in the S_3 subsite.

Based on the transition-state theory, the change in free energy between the free enzyme and substrate and the transition state ($\Delta\Delta G_T^{\ddagger}$) is equal to the binding energy released when the substrate is bound (ΔG_S) and the activation energy ($\Delta G^{\ddagger}$). $\Delta G^{\ddagger}$ represents the energy for the formation of the ES (enzyme-substrate) complex from the ground state to the transition state. This is shown in the following equation.

$$\Delta G_T^{\ddagger} = \Delta G_S + \Delta G^{\ddagger} = RT\ln(k_B T / h) - RT\ln(k_{cat} / K_m)$$

Under the assumption that $\Delta G^{\ddagger}$ is not significantly affected by mutations within the enzyme or changes in the pH of the assay, the differences in $\Delta\Delta G^{\ddagger}_{T}$ for mutant enzymes tested with different substrates can be evaluated as follows:

$$\Delta\Delta G^{\ddagger}_{T} = -RT\frac{\ln(k_{cat}/K_{m})_{mutant}}{(k_{cat}/K_{m})_{wildtype}}$$

The removal of the glutamic acid caused a relatively small reduction in the free energy of transition-state stabilization calculated by the equation based on the transition-state theory.

The average value of $\Delta\Delta G^{\ddagger}_{T}$ is +0.1 kcal/mole. Hence, mutating the Glu13 residue was not significant in terms of the energy contribution toward the specificity of the enzyme. No significant changes in the k_{cat}/K_{m} values were observed with the Glu287 mutants as well as the double mutant in all the P_3 substitutions tested. pH dependence studies with the arginine containing peptide were indicative of the titratable effect of the Glu13 residue. The energy change substituting the Glu287 residue with either a methionine or a glutamine averaged to +0.3 kcal/mole for the P_3 substitutions.

The S_2 Subsite

The Glu13Ala mutant enzyme exhibited a 3-fold increase in k_{cat}/K_{m} values for the leucine and the alanine substitutions in the P_2 position with respect to the wild type enzyme. As explained in Chapter 1, all the peptides bind to the active site in an extended β-strand conformation. Binding of a residue in one subsite can affect the binding in another subsite. Changing the Glu13 residue to an alanine has made the S_3 subsite generally hydrophobic. Hence, the alanine in the P_3 position of the peptide can position itself strongly in the hydrophobic environment. This strong anchoring in the S_3 subsite can effect the positioning of a leucine or an alanine residue in the S_2 subsite. The change in positioning of the peptide in the active site cleft of the mutant enzyme in comparison to the wild type enzyme may lead to the observed 3-fold increase in the k_{cat}/K_m values which reflect the catalytic efficiency. As discussed earlier, in Chapter 3 and in this chapter, the S_2 subsite of porcine pepsin is largely hydrophobic. Hence, the leucine and the alanine residue bind productively exhibiting the highest k_{cat}/K_m values with all the substitutions tested in the P_2 position.

Thus, the strong anchoring of the alanine residue in the S_3 subsite effects the interaction in the S_2 subsite. This observation suggests that there is an interdependency between the subsites of the enzyme. Detailed discussion of this observation is provided in Chapter 5.

In the case of the Glu287 mutants, a large effect on the k_{cat}/K_m parameter was observed. All the P_2 substitutions exhibited a 2-3 fold increase in k_{cat}/K_m value. This effect, as discussed earlier, was predominantly due to the increase in the k_{cat} number. By replacing the Glu287 residue with a methionine or a glutamine we have essentially made the enzyme more efficient toward cleavage of the peptides with various P_2 substitutions in comparison to the wild type enzyme. The contribution to the binding energy of the transition state with these replacements averages to -1.0 kcal/mole. This gain in interaction energy of the enzyme with the transition state is mainly due to the replacement of the polar glutamate to a less polar glutamine or a hydrophobic methionine residue.

These results suggest that the S_2/P_2 interaction is favorable with a hydrophobic residue in the S_2 subsite. A certain flexibility of the peptide may also be involved in the efficiency of catalysis. Binding of the substrates to the Glu287 mutants is enhanced by the, now, stronger hydrophobic environment of the S_2 subsite of the enzyme. This may lead to the decrease in the energy barrier of the transition state of the enzyme. This results in up to a 3- to 6-fold increase in k_{cat} values compared to the wild type enzyme.

pH dependence studies with the P_2 lysine containing peptide displayed a titratable response with a pK_a of around 4.3. This effect was interesting because the Glu287 residue, proposed to permit electrostatic interactions in the S_2 subsite, has now been replaced by non-titratable groups, namely, methionine and glutamine.

It is known that the strength of non-covalent interactions are dependent upon the nature of groups that are involved and the distance between them (Wells, 1990). The distance of charge-charge interactions is dependent upon 1/r, where r is the distance between the two groups involved in the charge-charge interaction (Fersht, 1977).

Pepsin has an abundance of negatively charged residues (15% of the amino acids of the protein). Some of these residues are surface exposed, while some are within the core of the enzyme in the vicinity of the active site cleft. The pH dependent effects observed for the Glu287 mutant enzymes may be attributed to the general electrostatic interactions with other negatively charged residues in the vicinity. At this juncture, it is difficult to pinpoint which residue may be involved in this electrostatic interaction. Two of the candidates are Glu297 and Asp290.

In conclusion, the S_3/P_3 interactions do not seem to make a major contribution toward the specificity of the enzyme with the exception of discrimination against positively charged residues in the P_3 position of the peptide. The dominant contributor to the enzymatic specificity seems to be the S_2/P_2 interactions. Also, the results reported in this chapter show an interdependency between the two subsites. Binding of a residue in one subsite can effect the binding requirements in another subsite. The specific energy contributions toward enzymatic specificity of the mutant enzymes will be elaborated in Chapter 5.

CHAPTER 5
COMPARISON OF KINETIC PARAMETERS OF PORCINE PEPSIN WITH CATHEPSIN D AND CATHEPSIN E

Introduction

Studies have been reported on the catalytic efficiency of pepsin and other members of the aspartic proteinase family of enzymes (Inouye and Fruton, 1967; Fruton, 1970, Sachdev and Fruton, 1970; Dunn et al , 1984, Dunn and Kay, 1985) The efficiency of pepsin catalysis has been shown to be dependent upon the length of the peptide substrate and nature of the amino acid side chains that make up the active site cleft initially referred to as "secondary" interactions by Fruton (Fruton, 1970; Voynick and Fruton, 1971). Such secondary interactions are important for enzymatic specificity (Dunn et al , 1987; Pohl and Dunn, 1988, Richards et al., 1990)

This chapter addresses the differences in these secondary interactions specifically in the S_2 and the S_3 subsites of three mammalian enzymes porcine pepsin: human cathepsin D and human cathepsin E.

This chapter also describes the energy contributions of the Glu13 and the Glu287 mutants toward enzymatic specificity

Results

The following results will specifically relate to the differences in the S_3 and S_2 subsite preferences for the mammalian enzymes, namely human cathepsin D and human cathepsin E, in comparison to porcine pepsin. Table 5-1 lists the comparative specificity constants (k_{cat}/K_m) of the three mammalian enzymes under consideration. The substrates were of the form Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu with systematic substitutions in the P_3 and P_2 positions.

S_3 Subsite

All substrates with substitutions in the P_3 position were cleaved efficiently by porcine pepsin with the exception of the P_3 arginine. Cathepsin D, on the other hand, exhibited a strong dislike for the substrates with all the substitutions tested. A similar trend was observed for cathepsin E, exhibiting low k_{cat}/K_m values, in comparison with porcine pepsin.

As explained earlier in Chapter 3, pH dependence studies carried out with a peptide with an arginine in the P_3 position indicated that the group that is titrated has a pK_a of around 4.2.

Table 5-1 Specificity constants (k_{cat}/K_m) for the enzyme-catalyzed cleavage of peptide substrates of the form Lys-Pro-*Xaa*-*Yaa*-Phe*Nph-Arg-Leu

PEPTIDES	PORCINE PEPSIN[b]	HUMAN CATHEPSIN D[c]	HUMAN CATHEPSIN E[d]
P_5-P_4-P_3 - P_2 - P_1*P_1'- P_2'- P_3'		k_{cat}/K_m ($\mu M^{-1}sec^{-1}$)	
Lys-Pro-Ala-Lys-Phe*Nph[a]-Arg-Leu	0.8	< 0.02	0.1
Lys-Pro-*Asp*-*Lys*-Phe*Nph-Arg-Leu	0.1	< 0.02	0.1
Lys-Pro-*Ser*-*Lys*-Phe*Nph-Arg-Leu	0.3	< 0.02	0.1
Lys-Pro-*Leu*-*Lys*-Phe*Nph-Arg-Leu	0.4	< 0.02	0.2
Lys-Pro-*Arg*-*Lys*-Phe*Nph-Arg-Leu	<0.02	< 0.02	0.1
Lys-Pro-*Ala*-*Ala*-Phe*Nph-Arg-Leu	0.8	0.2	0.7
Lys-Pro-*Ala*-*Asp*-Phe*Nph-Arg-Leu	0.5	0.3	0.1
Lys-Pro-*Ala*-*Ser*-Phe*Nph-Arg-Leu	0.5	0.1	0.7
Lys-Pro-*Ala*-*Leu*-Phe*Nph-Arg-Leu	0.7	1.0	2.0
Lys-Pro-*Ala*-*Arg*-Phe*Nph-Arg-Leu	0.5	<0.02	0.1

[a]Nph = *p*-nitrophenylalanine. Assay conditions for porcine pepsin were 0.1M sodium formate buffer, pH 3.5, 37°C. [c]cathepsin D, pH 3.7, 37°C (Scarborough, 1993) and for [d]cathepsin E, pH 4.5, 37°C.
Standard errors were within 15%.

This titratable effect with the arginine containing peptide was also seen with cathepsin E. Analyzing the crystal structure of porcine pepsin (Abad-Zapatero et al., 1991) and the rule-based molecular model of cathepsin E (unpublished data), one of the residues common to both the enzymes is glutamic acid 13 in the S_3 subsite

S_2 Subsite

Table 5-1 also represents the k_{cat}/K_m values exhibiting the S_2 subsite preferences of the three mammalian enzymes under consideration. The peptides were cleaved at measurable rates by porcine pepsin. The substitutions were not particularly discriminatory. Cathepsin D accommodates the hydrophobic leucine residue with the highest k_{cat}/K_m value for all the substitutions tested. Cathepsin E exhibited similar preferences for the P_2 substituents. The serine substitution was more productive in comparison to the cathepsin D results. The P_2 arginine containing peptide was not cleaved by cathepsin D.

Since all the peptides containing a positively charged residue in the P_2 (lysine or arginine) position were not cleaved by cathepsin D, a second generation of peptides was synthesized, to further explore the distinctions in the S_2 subsite (Scarborough, 1993; Scarborough et al., 1993; Scarborough and Dunn, 1994).

Table 5-2 Specificty constants (k_{cat}/K_m) of peptide substrates of the form Lys-Pro-Ile-Glu-Phe*Nph[a]-Arg-Leu with systematic substitutions in the P_2 position for porcine pepsin, human cathepsin D and human cathepsin E

								PORCINE PEPSIN[b]	CATHEPSIN D[c]	CATHEPSIN E[d]
P_5	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'		k_{cat}/K_m ($\mu M^{-1}s^{-1}$)	
Lys	Pro	Ile	*Leu*	Phe	Nph	Arg	Leu	1.4	2.0	2.0
Lys	Pro	Ile	*Ala*	Phe	Nph	Arg	Leu	1.2	0.8	3.0
Lys	Pro	Ile	*Nle*	Phe	Nph	Arg	Leu	1.0	ND[e]	7.0
Lys	Pro	Ile	*Glu*	Phe	Nph	Arg	Leu	1.0	0.9	7.0
Lys	Pro	Ile	*Ser*	Phe	Nph	Arg	Leu	0.9	0.4	4.0
Lys	Pro	Ile	*Asp*	Phe	Nph	Arg	Leu	0.9	0.3	4.0
Lys	Pro	Ile	*Arg*	Phe	Nph	Arg	Leu	0.8	0.1	1.0
Lys	Pro	Ile	*Ile*	Phe	Nph	Arg	Leu	0.8	0.9	8.0
Lys	Pro	Ile	*Val*	Phe	Nph	Arg	Leu	0.6	0.5	5.0
Lys	Pro	Ile	*Asn*	Phe	Nph	Arg	Leu	0.5	ND	4.0
Lys	Pro	Ile	*Gln*	Phe	Nph	Arg	Leu	0.3	0.2	3.0
Lys	Pro	Ile	*His*	Phe	Nph	Arg	Leu	0.2	NC	0.1
Lys	Pro	Ile	*Thr*	Phe	Nph	Arg	Leu	0.2	ND	2.0
Lys	Pro	Ile	*Lys*	Phe	Nph	Arg	Leu	0.2	NC[f]	0.4
Lys	Pro	Ile	*Pro*	Phe	Nph	Arg	Leu	NC	NC	NC

[a]Nph = *p*- nitrophenylalanine, assay conditions for [b]porcine pepsin were 0.1 M sodium formate buffer , pH 3.5, 37°C, [c]cathepsin D were 0.1 M sodium formate buffer, pH 3.7, 37°C (Scarborough, 1993) and for [d]cathepsin E were 0.1 M sodium formate buffer, pH 4.5, 37°C. [e]ND = Not determined, [f]NC = Not cleaved under standard assay conditions. Standard errors were within 15%.

These peptides were tested with porcine pepsin and cathepsin E to explore the distinctions in the S_2 subsite in detail of the three enzymes (Table 5-2).

The second generation of peptides has an isoleucine in the P_3 position and systematic substitutions in the P_2 position with fifteen different amino acids (Table 5-2). Isoleucine was chosen at the P_3 position because cathepsin D has a 4- to 5- fold preference for isoleucine in the P_3 position compared to alanine (Scarborough et al., 1993).

This series has generated excellent substrates for cathepsin D and cathepsin E. Increasing the hydrophobicity of the side chain in the P_2 position increased the catalytic efficiency for cathepsin D. In agreement with the observation in Table 5-1, the positively charged substitutions were not cleaved by cathepsin D.

For cathepsin E, the glutamic acid substitution exhibited approximately two-fold higher k_{cat}/K_m values than the aspartic acid. This can be ascribed to the extra methylene group in the glutamic acid residue, which can potentially participate more favorably with the predominantly hydrophobic S_2 pocket of the enzyme. This difference may also be related to the differences in the pK_a values of the side chains. Since the pK_a of the glutamate is slightly higher than the aspartate, it is possible that at pH 4.5 (assay pH for kinetic characterization of cathepsin E), the aspartate is completely deprotonated rendering it negatively charged, whereas the glutamic acid is partially deprotonated still exhibiting some hydrophobicity.

The arginine, histidine and lysine substitutions exhibited the lowest k_{cat}/K_m numbers for cathepsin E and cathepsin D. It can be concluded from this observation that like cathepsin D, cathepsin E also has a dislike for positively charged residues in the S_2 subsite.

Proline was totally unacceptable in the S_2 subsite for all the enzymes under study which may be explained due to the unique conformation of the proline side chain leading to unproductive binding.

Porcine pepsin also accommodates the hydrophobic substitutions readily due to the generally hydrophobic nature of the S_2 pocket. However, it also cleaves peptides with a positive charge in the P_2 position. This interaction may be possible due to the presence of a glutamic acid at position 287 which can be involved in electrostatic interactions with a positively charged residue in the P_2 position discussed in Chapters 3 and 4.

In general, considering the S_2 subsite of cathepsin D and cathepsin E, hydrophobic residues in the P_2 position of the substrate are optimal. Negatively charged and hydrogen bonding residues are acceptable and a strong dislike for positively charged residues is observed. Porcine pepsin also prefers hydrophobic residues largely. The difference, however, lies in the acceptance of positively charged residues in the P_2 position owing to the potential electrostatic interaction with Glu287.

Discussion

Comparing the crystal structures of the three mammalian enzymes being discussed, the S_3 subsite of all the enzymes is predominantly hydrophobic (Abad-Zapatero et al., 1991; Fusek et al., 1992, Metcalf and Fusek, 1993; Scarborough et al., 1993; unpublished data). As discussed earlier in Chapters 3 and 4, the S_3 subsite of porcine pepsin is defined by the residues Tyr75, Thr77, Phe111, Tyr114, Ala115, Pro116, Gly217, and Thr218 (Abad-Zapatero et al., 1991). The hydrophobic nature of the pocket can enhance hydrophobic interactions with the bulky β-branched isoleucine and the hydrophobic alanine residue to the same extent. The peptides may anchor strongly due to the size of the S_3 pocket in comparison to the size in cathepsin E or cathepsin D.

A significant difference which is seen in the S_3 subsite of cathepsin D is that Thr111 and Ala114 replace Phe111 and Tyr114 residues, respectively, in porcine pepsin. This may result in an increase in the size of the S_3 pocket of cathepsin D. Ile at P_3 of the substrate will be optimal for packing in the S_3 pocket leading to tighter binding. This observation may be explained by the size of the S_3 subsite, as well as the nature of the amino acids that define the pocket. It should be emphasized that the peptides bind in the active site of aspartic proteinases in an extended fashion and there is an interdependency between the subsites. Based on the model of cathepsin E, the S_3 pocket is large in comparison to that of pepsin and is essentially composed of polar amino acids. Also, the Phe111 is replaced by a Tyr111.

Comparing the S_2 subsite in crystal structures of cathepsin D, porcine pepsin and the rule-based model of cathepsin E, it can be seen that position 287 can greatly influence the binding of the substrates with a positive charge in the P_2 position. Glu287 in porcine pepsin which borders the S_2 and S_4 subsite can be involved in electrostatic interactions with a positively charged lysine residue. Cathepsin D and cathepsin E lack this interaction due to the presence of hydrophobic methionine and uncharged glutamine at position 287, respectively. Hence, the presence of the negative charge in porcine pepsin may influence greatly the ability to cleave peptides with a positive charge in the P_2 position. The S_2 subsite of cathepsin D is largely hydrophobic, promoting strong interactions with hydrophobic residues, such as leucine. The negatively charged aspartate and hydrogen bonding serine are moderately acceptable in this predominantly hydrophobic environment. Supporting the earlier data, positively charged arginine is unacceptable in the S_2 subsite. In cathepsin E the acceptance of a serine residue in the P_2 position can be supported by the presence of three threonine residues namely, Thr77, Thr218, Thr222 which define the S_2 binding pocket in human cathepsin E. The charged aspartate and arginine were least acceptable This may be attributed to the lack of the glutamic acid in the S_2 subsite which is present in porcine pepsin and also the possibility of the number of rotatable bonds of the positively charged side-chains of lysine or arginine that may control the efficiency of catalysis

Enzymatic catalysis is controlled by the binding of the substrate and the transition state to the active site cleft of the enzyme. The binding of the substrate to the enzyme is controlled by a variety of forces which include hydrogen bonding, electrostatic, hydrophobic and van der Waals interactions, The three-dimensional crystal structure in combination with detailed kinetic analyses can allow us to predict the roles of residues that may play a prominent role in catalysis, These analyses in combination with site directed mutagenesis permits the exploration of the effects of the critical residues in enzymatic specificity and catalysis.

Combining the three approaches, I have attempted to evaluate the role of two glutamic acid residues, namely, Glu13 in the S_3 subsite and Glu287 in the S_2 subsite of porcine pepsin. The replacements at these two positions with residues corresponding in cathepsin D or cathepsin E, two mammalian enzymes of the same family, have provided insights on the importance of the S_2/P_2 interactions. The data in Chapter 3 and Chapter 4 are suggestive of the interdependency between the two subsites where the alterations have been made. This chapter will discuss the contributions of the Glu13Gln, Glu287Met and the Glu13/Glu287 mutants toward the energy of binding of the transition state tested with peptides with substitutions in the P_3 and P_2 positions.

Change in free energy caused by a mutation is expressed relative to the wild type protein as $\Delta\Delta G_{(X)}$ where X = a single mutation. The free energy change for the two mutations X and Y can be related to the energy of the double mutant by the following equation (Carter et al., 1984; Ackers and Smith, 1985; Wells, 1990):

$$\Delta\Delta G_{(X,Y)} = \Delta\Delta G_{(X)} + \Delta\Delta G_{(Y)} + \Delta G_{I}$$

The ΔG_{I} term is known as the coupling energy (Carter et al., 1984). This term defines the change in energy of interaction between the two sites that have been mutated. This value is either positive or negative depending on whether there is an increase or decrease in the functional property of the enzyme (Wells, 1990).

When the mutations are independent of each other, the coupling energy is negligible and the free energy change of the double mutant is equal to the sum of the free energies of the two single mutants. This is shown in the following equation:

$$\Delta\Delta G_{(\mathrm{Double})} = \Delta\Delta G_{(\mathrm{Glu13Gln})} + \Delta\Delta G_{(\mathrm{Glu287Met})}$$

Such an interaction is termed as "simple additivity". This effect is clearly observed with the mutants tested with the P_3 substituents of the peptide of the form Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu. The overall change in free energy going from the wild type to the double mutant averages to about +0.5 kcal/mole, which is the sum of the energies of the individual mutations for a given peptide. Hence, these mutations are perfectly additive with these set of peptides calculated using the following equation (Wilkinson et al., 1983). The results are shown in Table 5-3.

$$\Delta\Delta G_T^{\ddagger} = -RT\,\frac{\ln(k_{cat}/K_m)_{Glu13Gln}}{(k_{cat}/K_m)_{wildtype}}$$

If the interactions are perfectly additive, then the following relation holds true:

$\Delta G_1 = \Delta G_2$ and $\Delta G_3 = \Delta G_4$ (Carter et al., 1984)

where, ΔG_1 change in free energy caused by the Glu13Gln mutation expressed relative to the wild type protein, ΔG_2 = change in free energy caused by the double mutation expressed relative to Glu287Met protein, ΔG_3 = change in free energy caused by the Glu287Met mutation a relative to the wild type protein, and ΔG_4 = change in free energy caused by the double mutation expressed relative to the Glu13Gln mutant protein. However, if the two mutations were interacting in some way with each other or if the mutations give rise to local conformational changes in the enzyme or the enzyme-substrate complex, the changes in the interaction energies of the single mutants do not add up, i.e.,

$$\Delta G_1 \neq \Delta G_2 \text{ and } \Delta G_3 \neq \Delta G_4, \text{ and}$$
$$(\Delta G_1 - \Delta G_2) = (\Delta G_3 - \Delta G_4) \neq 0$$ (Carter et al., 1984).

The change in $| \Delta G_1 - \Delta G_2 |$ is defined as the coupling energy between the two sites that have been mutated. The coupling energy in my case measures the energy difference measured between the double mutant (Glu13Gln/Glu287Met) and the two single mutants measured separately (Glu13Gln, Glu287Met). This phenomenon is termed as "non-additivity" (Wells, 1990). The non additivity effect can exist when the mutations at the two different positions change the interaction between the two subsites. As explained earlier in Chapter 4, there is a strong interdependency between the S_2 and S_3 subsites. These interactions can take place via electrostatic attractions, hydrogen bonding, hydrophobic interactions, and even through conformational changes within the peptide. Such changes can also be a result of replacement with a chemically different side chain.

Such deviations from simple additivity is reflected in the energy contributions of the double mutant cycles with the peptides of the form Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu systematically substituted with residues in the P_2 position.

The results are outlined in Table 5-4. There is a marked preference for all the P_2 peptide substitutions tested for the Glu287Met and Glu287Gln mutants (Table 4.2 in Chapter 4). Based on the energy contributions outlined in Table 5-4, the results are indicative of a favorable average transitional state binding energy of -1.0 kcal/mole. This suggests that the orientation of the P_2 side chain in the Glu287 mutants is different leading to productive binding and higher enzymatic specificity toward these peptides. The energy terms are not simply additive with these set of peptides. Hence, changing the residue in the S_2 subsite has an effect on catalysis. This effect was not significant with the same set of peptides and the Glu13 mutants. In thermodynamic terms for the non-additivity effect, $\Delta G_1 \neq \Delta G_2$ and $\Delta G_3 \neq \Delta G_4$. The increase in catalytic activity by the mutant enzyme toward the same substrate clearly implies that the structure of the mutants is perfectly normal.

This suggests that the ES interface may exhibit some flexibility in the S_2/P_2 region. This flexibility may also arise due to movement of the side chain of the substrate. Similar results on non-additivity have been extensively reported on the serine proteinase subtilisin Savinase from *Bacillus lentus* (Bech et al., 1993).

Table 5-3. Analysis of the double mutant cycles of porcine pepsin with the cleavage of the peptide of the form Lys-Pro-Ala-Lys-Phe*Nph[a]-Arg-Leu with substitutions in the P_3 position

Peptide	pH 3.5		
Lys-Pro-Ala-Lys-Phe*Nph-Arg-Leu	800,000	*0.2*	700,000
	WT[b]		E287M[d]
	0.4		*0.4*
	E13Q[c]		DBL[e]
	400,000	*0.2*	300,000
Lys-Pro-Ser-Lys-Phe*Nph-Arg-Leu	300,000	*0.2*	200,000
	WT		E287M
	0		*0.4*
	E13Q		DBL
	300,000	*0.6*	100,000
Lys-Pro-Asp-Lys-Phe*Nph-Arg-Leu	200,000	*0.4*	100,000
	WT		E287M
	0		0
	Glu13Gln		DBL
	200,000	*0.4*	100,000
Lys-Pro-Leu-Lys-Phe*Nph-Arg-Leu	600,000	*0.2*	400,000
	WT		E287M
	0.4		*0.4*
	E13Q		DBL
	300,000	*0.2*	200,000

[a] Nph - *p* - nitrophenylalanine, *italics* = ΔΔG (kcal mol^{-1}), numbers above and below enzyme type are k_{cat}/K_m ($M^{-1}s^{-1}$), [b] WT = Wild type enzyme, [c] E13Q = Glu13Gln mutant enzyme, [d] E287M =Glu287Met mutant enzyme, [e] DBL = double mutant (Glu13Gln/Glu287Met)

Table 5-4 Analysis of the double mutant cycles of porcine pepsin with the cleavage of the peptide of the form Lys-Pro-Ala-Lys-Phe*Nph^a-Arg-Leu with substitutions in the P_2 position

Peptide	pH 3.5		
Lys-Pro-Ala-Ala-Phe*Nph-Arg-Leu	800,000	*-1.0*	4000,000
	WT[b]		E287M[d]
	0.1		*0.9*
	E13Q[c]		DBL[e]
	700,000	*-0.2*	1000,000
Lys-Pro-Ala-Arg-Phe*Nph-Arg-Leu	500,000	*-0.9*	2000,000
	WT		E287M
	-0.1		*1.2*
	E13Q		DBL
	600,000	*0.4*	300,000
Lys-Pro-Ala-Leu-Phe*Nph-Arg-Leu	700,000	*-0.6*	2000,000
	WT		E287M
	0.5		*1.2*
	E13Q		DBL
	300,000	*0*	300,000
Lys-Pro-Ala-Asp-Phe*Nph-Arg-Leu	500 000	*-1.0*	3000 000
	WT		E287M
	0.3		*0.9*
	E13Q		DBL
	300,000	*-0.5*	700,000
Lys-Pro-Ala-Ser-Phe*Nph-Arg-Leu	500,000	*-0.9*	2000 000
	WT		E287M
	1.0		*0.9*
	E13Q		DBL
	100,000	*-1.0*	500,000

[a] Nph = *p*-nitrophenylalanine, *italics* = ΔΔG (kcal mol^{-1}), numbers above and below enzyme type are k_{cat}/K_m ($M^{-1}s^{-1}$), [b] WT = Wild type enzyme, [c] E13Q = Glu13Gln mutant enzyme, [d] E287M =Glu287Met mutant enzyme, [e] DBL = double mutant (Glu13Gln/Glu287Met)

The lysine substitution did not effect the specificity constants with the Glu287 mutants. This observations may imply that the geometry and the number of rotatable bonds of the positively charged residue may exert an effect on the enzymatic specificity. This effect was clearly observed even with the cathepsin D and cathepsin E results tested with the same set of peptides (refer to Table 5-1 and 5-2). Mutagenesis experiments carried out with cathepsin D, where the Met287 residue in the S_2 subsite, was replaced by a glutamic acid to mimic porcine pepsin was supportive of this interchangeable specificity to some extent. Cathepsin D, which has a strong dislike for positively charged residues in the P_2 position, exhibited measurable cleavage of such peptides in mutants replacing the Met287 with a glutamic acid (Scarborough, 1993; Scarborough and Dunn, 1994). In my case with porcine pepsin, by making the S_2 subsite generally hydrophobic by replacing the polar glutamate with a methionine, the enzyme now exhibits greater specificity toward hydrophobic substitutions similar to cathepsin D.

Reports on structural changes induced upon the binding of an inhibitor or a substrate to the active site of pepsin have been reported in the literature (Abad-Zapatero et al., 1991). There exists a flexible sub-domain within the carboxyl lobe of pepsin and the fungal aspartic proteinases (Suguna et al., 1977; James and Sielecki, 1983; Abad-Zapatero et al., 1990, 1991; Blundell et al., 1990, Gilliland et al., 1990). The residues that are involved in this flexible sub-domain are 196-209, 221-238, 245-277 and 282-291. Position 287 falls within this mobile region. It may be possible that the removal of the Glu287 residue may reduce some steric constraints within the interface of the S_2/P_2 region. This release in constraints may promote the binding of the P_2 residue in the S_2 subsite.

All the data presented here suggest that there is an interdependency between the S_3 and S_2 subsites. Not so favorable interactions at one subsite may promote favorable interactions in the other as observed with the alanine and leucine containing peptides at the P_2 position with the Glu13 mutants. The S_2-P_2 interaction may cause a rearrangement of the peptides to decrease the energy barrier to the transition state, increasing the enzymatic specificity of the enzyme.

In conclusion, I have attempted to mimic some of the S_2 specificity of cathepsin D or cathepsin E into porcine pepsin. Based on all our kinetic data, the changing of Glu287 to methionine or glutamine was favorable with the hydrophobic P_2 substituted peptides tested. Although the active site remained intact, the changing of a single amino acid residue in the S_2 subsite had a major influence on catalysis, with selectivities similar to cathepsin D and cathepsin E. These mutations have proven useful in evaluating the contribution of the S_2 subsite toward enzymatic specificity of porcine pepsin.

CHAPTER 6
CONCLUSIONS AND FUTURE DIRECTIONS

Protein engineering has become an important tool to study structure-function relationship of proteins. Using a combined approach of three-dimensional structural information, site directed mutagenesis and kinetic analyses, the work presented here details the roles of the S_2 and S_3 subsites of porcine pepsin in enzymatic specificity.

The results have revealed the importance of the S_2/P_2 interactions involving the Glu287 residue in catalysis. pH dependence studies carried out with the peptide containing a lysine residue in the P_2 position, with the Glu287 mutants (Chapter 4), are suggestive of other potential residues in the vicinity of the S_2 subsite which may be involved in general electrostatic interactions. Two of the candidates are Glu297 and Asp290 which are 8-10 Å away from the lysine side chain in the P_2 position of the peptide. Triple mutants of Glu287, Asp290 and Glu297 may reveal the possibilities of these electrostatic interactions.

As discussed earlier in Chapter 4, pepsin has an abundance of negative charged residues in comparison to the other members of the family. The negative charges on the surface of a protein can affect structure, stability as well as the biological function of the protein.

These charge-charge interactions may be local involving residues Glu287 and the lysine in the P_2 position of the peptide or between the Glu13 residue and an arginine in the P_3 position of the substrate. The effect of changing solvent exposed residues to see the effect on catalysis and the pH dependent kinetic parameters will be of interest.

Examination of the specificity of chimeric proteins will be another aspect that can be explored. An example would be to make a chimeric protein with the C-terminal domain of cathepsin E or cathepsin D and the N-terminal domain of porcine pepsin. This can be followed by making site directed mutants at position 287 and test the pH dependent activity and effects of S_2/P_2 interactions. This study may give us some insights into the general electrostatic interactions taking place in porcine pepsin. As explained in Chapter 5, pepsin has an abundance of negatively charged residues compared to the other members of this family. Based on the crystal structure of the enzyme the negatively charged residues are in the vicinity of the active site toward the C-terminal domain. Hence these chimeric proteins will aids in understanding the general electrostatic interactions involving the negatively charged residues in the C-terminal domain of porcine pepsin.

The results in this thesis have also provided us understanding about general protein-ligand interactions and the contribution of hydrophobic and electrostatic interactions toward enzymatic specificity.

The detailed specific interactions taking place in the subsites of the enzymes are very important in the development of therapeutic agents in the pharmaceutical industry. This involves blocking the target enzyme without attenuating the normal role of the other enzymes in this family.

The subsites in the prime region remain to be explored. As described in Chapter 3, the S2' subsite has a preference for a positively charged residue with all the substitutions tested. The contribution of this specificity can be examined using site directed mutagenesis followed by kinetic analyses.

LIST OF REFERENCES

Abad-Zapatero, C., T. J. Rydel and J. Erickson (1990) *Proteins. Struct. Funct. Genet.* 8, 62-81.

Abad-Zapatero, C., T. J. Rydel, D. J. Neidhart, J. Luly and J. W. Erickson (1991). in *Structure and Function of Aspartic Proteinases: Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 9-21. Plenum Press, New York.

Ackers, G. K. and F. R. Smith (1985). *Annu. Rev. Biochem.* 54, 597-629.

Aikawa, J., R. Hiramatsu, M. Nishiyama, S. Horinouchi and T. Beppu (1991). in *Structure and Function of Aspartic Proteinases: Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 233-242. Plenum Press, New York.

Al-Janabi, A., J. A. Hartsuck and J. Tang (1972). *J. Biol. Chem.* 247, 4628-4632.

Andreeva, N. S., A. S. Zdanov, A. E. Gustchina and A. A. Fedorov (1983), *J. Biol. Chem.* 259, 11353-11365.

Barrett, A. J. (1975). in *Proteases and Biological Control* (eds. E. Reich, D. B. Rifkin and E. Shaw) 467 480. Cold Spring Harbor Lab., Cold Spring Harbor, New York.

Barrett, A. J. and H. Kirschke (1981). *Methods in Enzymology* 80, 535-581.

Barrett, A. J. (1980). in *Mammalian Proteases: A Glossary and Bibliography, Vol I: Endopeptidases* (ed. J. K. McDonald) 303-356. Academic Press, New York.

Bech, L. M., S. B. Sorensen and K. Breddam (1992). *Eur. J. Biochem.* 209, 869-874.

Bech, L. M., S. B. Sorenson and K. Breddam (1993). *Biochemistry 32*, 2845-2852.

Berger, A. and I. Schechter (1970). *Phil. Trans. Roy. Soc. 257*, 249-264.

Blum, M., A. Cunningham, M. Bendiner and T. Hofmann (1985). *Biochem. Soc. Trans. 13*, 1044-1046.

Blundell, T. L., J. Cooper, S. I. Foundling, D. M. Jones, B. Attresh and M. Szelke (1987). *Biochemistry 26*, 5585-5590.

Blundell, T. L., J. A. Jenkins, B. T. Sewell, L. H. Pearl, J. B. Cooper, I. J. Tickle, B. Veerapandian and S. P. Wood (1990). *J. Mol. Biol. 211*, 919-41.

Bode, W., E. Papamokos and D. Musil (1987). *Eur. J. Biochem. 166*, 673-692.

Bott, R., E. Subramanian and D. R. Davies (1982). *Biochemistry 21*, 6956-6962.

Brouillet, J-P., C. Theillet, T. Maudelonde, A. Defrenne, J. Simony-Lafonteine, J. Sertour, H. Pujol, F. Jeanteur and H. Rochefort (1990). *Eur. J. Cancer 26*, 437-441.

Bryan, P., M. W. Pantoliano, S. G. Quill, H. Y. Hsiao and T. Poulos (1986). *Proc. Natl. Acad. Sci. U. S. A. 83*, 3743-3745.

Carter, P. J., G. Winter, A. J. Wilkinson and A. R. Fersht (1984). *Cell 38*, 835-840.

Cooper, J., S. Foundling, A. Hemmings, T. L. Blundell, D. M. Jones, A. Hallett and M. Szelke (1987). *Eur. Jour. Biochem. 169*, 215-221.

Corey, D. R. and C. S. Craik (1992). *J. Am. Chem. Soc. 114*, 1784-1790.

Craik, C., C. Largman, T. Fletcher, S. Roczniak, P. Barr, R. Fletterick and W. Rutter (1985). *Science 228*, 291-296.

Craik, C , S Roczniak, C Largman and W Rutter (1987) *Science 237*, 909-913

Dunn, B M (1992) In *Mechanisms of Biological Importance* (ed. J M Coxon) 213-241, Jai Press Ltd , Greenwich, Connecticut

Dunn, B M , M Jimenez, B. F Parten, B P. Valler, C E Rolph and J Kay (1986) *Biochem J. 237*, 899-906,

Dunn, B. M , B. Kammermann and K R, McCurry (1984). *Anal. Biochem 138*, 68-73

Dunn, B M and J. Kay (1985). Biochem *Soc Trans 13*, 1041-1043.

Dunn, B M , M J Valler, C E Rolph, S, I Foundling, M Jimenez and J. Kay (1987) *Biochem Biophys. Acta. 913*, 122-130

Estell, D. A , T. P Graycar, J. V Miller, D B. Powers, J P. Burnier, P. G. Ng and J A. Wells (1986) *Science 233*, 655-663.

Estell, D , T Graycar and J. Wells (1985) *J. Biol Chem 260*, 6518-6521.

Fersht, A (1977) *Enzyme Structure and Mechanism* W H Freeman and Company, San Francisco

Foltmann, B (1981) *Essays Biochem. 17*, 52-84

Fruton, J S (1970) *Adv Enzymol. Related Areas Mol Biol. 33*, 401-443.

Fusek, M., M. Baudys and F Metcalf (1992) *J. Mol Biol. 226*, 555-557,

Gilliland, G L , E L. Winborne, J. Wachman and A. Wlodawer (1990), *Proteins: Struct Funct Genet* 8, 82-101

Goldberg, G. I., S. M. Wilhelm, A. Kronberger, E. A. Bauer, G. A. Grant and A. Z. Eisen (1986). *J. Biol. Chem.* *261*, 6600-6605

Graves, M. C. (1991) in *Structure and Function of the Aspartic Proteinases. Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 395-405 Plenum Press, New York.

Greenlee, W. J. (1987). *Pharm. Res.* *4*, 364-374.

Gron, H., M. Meldal and K. Breddam (1992). *Biochemistry* *31*, 6011-6018

Hasty, K. A., T. F. Pourmotabbed, G. I. Goldberg, J. P. Thompson, D. G. Spinella, R. M. Stevens and C. L. Mainardi (1990). *J. Biol. Chem,* *265*, 11421-11424.

Henderson, P. F. (1972) *Biochem. J.* *127*, 321-333.

Herriott, R. M. (1939). *J. of Gen. Phys.* *22*, 65-78.

Ho, S. N., H. D. Hunt, R. M. Horton, Pullen J. K. and L. R. Pease (1989) *Gene* 77, 51-59

Hofmann, T., B. Allen, M. Bendiner, M. Blum and A. Cunningham (1988). *Biochemistry* *27*, 1140-1146.

Hofmann, T., R. Hodges and M. N. G. James (1984). *Biochemistry* *23*, 635-643.

Inouye, K. and J. Fruton (1967). *Biochemistry* *6*, 1765-1777.

James, M. N. G. and A. Sielecki (1983) *J. Mol. Biol.* *163*, 299-361.

James, M. N. G. and A. Sielecki (1986) *Nature* *319*, 33-38

Jupp, R. A., B. M. Dunn, J. W. Jacobs, G. Vlasuk, K. E. Arcuri, D. F. Veber, D. S. Perlow, L. S. Payne, J. Boger, S. deLaszlo, F. K. Chakravarty, J. tenBroeke, D. G. Hangauer, D. Ondeyka, W. J. Greenlee and J. Kay (1990). *Biochem. J. 265*, 871-878.

Kamphuis, I. G., J. Drenth and E. N. Baker (1985). *J. Mol. Biol. 182*, 317-329.

Kamphuis, I. G., K. H. Kalk, M. B. A. Swarts and J. Drenth (1984). *J. Mol. Biol. 179*, 233-256.

Kay, J. and B. M. Dunn (1992). *Scand. J. Clin. Lab. Invest. 52*, 23-30.

Kho, C. and H. Zarbl (1992). *Biotechniques 12*, 228-230.

Khouri, H. E., C. Plouffe, S. Hasnain, T. Hirama, A. C. Storer and R. Menard (1991). *Biochem. J. 275*, 751-757.

Kraut, J. (1977). *Ann. Rev. Biochem. 46*, 331-359.

Kuzmic, P., C-Q. Sun, Z-C. Zhao and D. H. Rich (1991). in *Structure and Function of Aspartic Proteinases: Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 75-86. Plenum Press, New York.

Lees, W. E., S. Kalinka, J. Meech, S. J. Capper, N. D. Cook and J. Kay (1990). *FEBS Lett. 273*, 99-102.

Lin, X. L., M. Fusek, Z. Chen, G. Koelsh, H. F. Han, J. A. Hartsuck and J. Tang (1991). in *Structure and Function of Aspartic Proteinases: Genetics, Structures* and *Mechanisms* (ed. B. M. Dunn) 1-8. Plenum Press, New York.

Lowther, W. T., Z. Chen, X. L. Lin, J. Tang and B. M. Dunn (1992). in *Structure and Function of Aspartic Proteinases: Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 275-279. Plenum Press, New York.

Lyons, J. G., B. Birkedal-Hansen, W. G. I. Moore, R. L. O'Grady and H. Birkedal-Hansen (1991). *Biochemistry 30*, 1449-1456.

Mallya, S. K., K. A. Mookhtiar, Y. Gao, R. Brew, M. Dioszegi, B. Birkedal-Hansen and H. E. Van Wart (1990). *Biochemistry 29*, 10628-10634.

Maniatis, T., E. F. Fritsch and J. Sambrook (1982). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Lab., Cold Spring Harbor, New York.

Marciniszyn, J., J. S. Huang, J. A. Hartsuck and J. Tang (1976). *J. Biol. Chem. 251*, 7095-7102

McPhalen, C. A. and M. N. G. James (1988). *Biochemistry 27*, 6582-6598.

Mcphie, P. (1972). *J. Biol. Chem. 247*, 4277-4281

Menard, R., H. Khouri, C. Plouffe, R. Dupras, D. Ripoll, T. Vernet, D. Tessier, F. Laliberte, D. Thomas and A. Storer (1990). *Biochemistry 29*, 6706-6713.

Menard, R., H. E. Khouri, C. Plouffe, P. Laflamme, R. Dupras, T. Vernet, D. C. Tessier, D. Y. Thomas and A. C. Storer (1991). *Biochemistry 30*, 5531-5538.

Metcalf, P. and M. Fusek (1993). *EMBO J. 12*, 1293-1302.

Meyer, S. L. (1975). *Data Analysis for Scientists and Engineers.*, John Wiley and Sons, Inc., New York.

Mookhtiar, K. A. and H. E. van Wart (1990). *Biochemistry 29*, 10620-10627

Muller, D., B. Quantin, M. C. Gesnel, R. Millon-Collard, J. Abecassis and R. Breatnach (1988). *Biochem. J. 253*, 187-192.

Ong, E. and G. Perlmann (1968). *J. Biol. Chem. 243*:23, 6104-6109

Parris, K. D., D. J. Hoover and D. R. Davies (1991) in *Structure and Function of Aspartic Proteinases: Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 217-231 Plenum Press, New York.

Pearson, J. P., R. Ward, A. Allen, N. B. Roberts and W. H. Taylor (1986) *Gut 27*, 243-248.

Pedersen, V. and B. Foltmann (1973). *FEBS. Lett. 35*, 255-256.

Perona, J. J., C. A. Tsu, M. E. Mcgrath, C. S. Craik and R. J. Pletterick (1993) *J. Mol. Biol. 230*, 934-949.

Pillai, S. and J. Zull (1986). *J. Biol. Chem. 261.32*, 14919-14923.

Pohl, J. and B. M. Dunn (1988) *Biochemistry 27*, 4827-4834.

Powers, J. C., A. D. Harley and D. V. Myers (1977). *Adv. Exp. Med. Biol. 95*, 141-157.

Rao, C. M., P. E. Scarborough, W. T. Lowther, J. Kay, B. Batley, S. Rapundalo, S. Klutchko, M. D. Taylor and B. M. Dunn (1991). in *Structure and Function of Aspartic Proteinases: Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 143-147. Plenum Press, New York.

Alchards, A. D., L. H. Phylip, W. G. Farmerie, P. E. Scarborough, A. Alvarez, B. M. Dunn, P. H. Hirel, J. Konvalinka, P. Strop, L. Pavlickova, V. Kostka and J. Kay (1990) *J. Biol. Chem. 265*, 7733-7736.

Sachdev, G. and J. Fruton (1970). *Biochemistry 9*, 4465-4470.

Sakai, H., T. Saku, Y. Kato and K. Yamamoto (1989). *Biochem. Biophys. Acta 991*, 367-375.

Sali, A., B. Veerapandian, J. B. Cooper, S. I. Foundling, D. J. Hoover and T. L. Blundell (1989). *EMBO J. 8*, 2179-88.

Sali, A., B. Veerapandian, J. B. Cooper, D. S. Moss, T. Hofmann and T. L. Blundell (1992) *Proteins Struct. Funct. Genet.* *12*, 158-170.

Salo, T., T. Turpeenniemi-Hujanen and K. Tryggvason (1985) *J. Biol. Chem.* *260*, 8526-8531.

Sanger, F., S. Nicklen and A. R. Coulson (1977). *Proc. Natl. Acad. Sci. U.S.A.* *74*, 5463-5467.

Scarborough, P. E. (1993) Unique Active Site Specificity of Human Cathepsin D as Studied Through Kinetics, Molecular Modeling and Residue-Specific Mutagenesis. Ph. D. Dissertation, University of Florida, Gainesville.

Scarborough, P. E. and B. M. Dunn (1994). *Protein Eng.* *7*, 495-502.

Scarborough, P. E., K. Guruprasad, C. Topham, G. R. Richo, G. E. Conner, T. L. Blundell and B. M. Dunn (1993). *Protein Sci.* *2*, 264-276.

Scarborough, P. E., G. R. Richo, J. Kay, G. E. Conner and B. M. Dunn (1991). in *Structure and Function of the Aspartic Proteinases: Genetics, Structures, and Mechanisms* (ed. B. M. Dunn) 343-347. Plenum Press, New York.

Schagger, H. and G. von Jagow (1987) *Anal. Biochem.* *166*, 368-379.

Schechter, I. and A. Berger (1967). Biochem. Biophys. *Res. Comm.* *27*, 157-162.

Sepulveda, P., J. Marciniszyn, D. Liu and J. Tang (1975). J. Biol. *Chem.* *250*, 5082-5088.

Sielecki, A., A. Fedorov, A. Boodhoo, N. Andreeva and M. N. G. James (1990) *J. Mol. Biol.* *214*, 143-170.

Stricklin, G. P., E. A. Bauer, J. J. Jeffrey and A. Z. Eisen (1977) Biochemistry *16*, 1607-1615.

Strop, P., J. Sedlacek, J. Stys, Z. Kaderabkova, I. Bleha, L. Pavlickova, J. Pohl, M. Fabry, V. Kostka, M. Newman, C. Frazao, A. Shearer, I. J. Tickle and T. L. Blundell (1990). *Biochemistry 29*, 9863-9871.

Studier, W. F., A. H. Rosenberg, J. J. Dunn and J. W. Dubendorff (1990). *Methods Enzymol. 185*, 60-89.

Suguna, K., E. A. Padlan, C. W. Smith, W. D. Carlson and D. R. Davies (1977). *Proc. Natl. Acad. Sci. U.S.A. 74*, 556-559.

Suguna, K., R. R. Bott, E. A. Padlan, E. Subramanian, S. Sheriff, G. H. Cohen and D. R. Davies (1987a). *J. Mol. Biol. 196*, 877-900.

Suguna, K., E. Padlan, C. Smith, W. Carlson and D. Davies (1987b). *Proc. Natl. Acad. Sci. U.S.A. 84*, 7009-7013.

Takeuchi, Y., S. Noguchi, Y. Satow, S. Kojima, I. Kumagai, K. Miura, K. T. Nakamura and Y. Mitsui (1991a). *Protein Eng. 4*, 501-508.

Takeuchi, Y., Y. Satow, K. T. Nakamura and Y. Mitsui (1991b). *J. Mol. Biol. 221*, 309-325.

Thomas, P., A. Russell and A. Fersht (1985). *Nature 318*, 375-376.

Umezawa, T., H. Aoyagi, H. Morishima, M. Matsuzaki, M. Hamada and T. Takeuchi (1970). *J. Antibiotics 23*, 259-262.

Vallee, B. L., A. Galdes, D. S. Auld and J. F. Riordan (1983). in *Zinc Enzymes* (ed. T. G. Spiro) 25-75. Wiley, New York.

Veerapandian, B., J. B. Cooper, A. Sali and T. L. Blundell (1990). *J. Mol. Biol. 216*, 1017-1029.

Voynick, I. M. and J. S. Fruton (1971). *Proc. Natl. Acad. Sci. U.S.A. 68*, 257-259.

Walker, V. and W. H. Taylor (1990). *Gut 21*, 766-771.

Wells, J. A. (1990). *Biochemistry 29*, 9509-9517.

Wells, J. A. and D. B. Powers (1996). *J. Biol. Chem. 261*, 6564-6570.

Wells, J. A., D. B. Powers, R. R. Bott, T. P. Graycar and D. A. Estell (1997). *Proc. Natl. Acad. Sci. U.S.A.* 94, 1219-1223.

Wilkinson, A. J., A. R. Fersht, D. M. Blow and G. Winter (1993). *Biochemistry 22*, 3581-3586.

Woessner, J. F. J. (1991). *FASEB J. 5*, 2145-2154.

Yamamoto, K., S. Sakai, E. Ueno and Y. Kato (1991). in *Structure and Function of Aspartic Proteinases: Genetics, Structures and Mechanisms* (ed. B. M. Dunn) 297-306. Plenum Press, New York

BIOGRAPHICAL SKETCH

Chetana Manohar Rao was born on June 4, 1963, in Bombay, India. Chetana received the degree of Bachelor of Science, majoring in biochemistry and botany in 1983, from the University of Bombay. She was awarded a gold medal for placing first in the University in both her major subjects. In 1985, she completed her Master of Science degree in biochemistry from the University of Bombay. In the next two years she pursued her research career as a research assistant at The Boots Pharmaceutical Company in Bombay. Her work focused on the pharmacokinetic aspects of drug delivery in living systems. She came to the United States of America in 1988. Her keen interests in biochemistry and molecular biology brought her to the University of Florida where she began her doctoral studies under the mentorship of Prof. Ben M. Dunn. After graduation, Chetana will continue her postdoctoral work at the University of Florida in Prof. Ben Dunn's laboratory.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ben M. Dunn

Ben M. Dunn, Chairman
Professor of Biochemistry
and Molecular Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Brian Cain
Assistant Professor of
Biochemistry and Molecular
Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Harry Nick

Harry Nick
Associate Professor of
Biochemistry and Molecular
Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

David Silverman
Professor of Pharmacology
and Therapeutics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

D. E. Richardson

David Richardson
Associate Professor of
Chemistry

This dissertation was submitted to the Graduate Faculty of the College of Medicine and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1994

Robert T Watson for Allen Neims

Dean, College of Medicine

Karen A Holbrook

Dean, Graduate School

UNIVERSITY OF FLORIDA
3 1262 08554 8187